22

PS/G2

THE BRITISH JOURNAL OF STATISTICAL PSYCHOLOGY

EDITED BY

CYRIL BURT AND JOHN WHITFIELD

WITH THE ASSISTANCE OF

CHARLOTTE BANKS

AND THE FOLLOWING EDITORIAL BOARD

A. C. AITKEN
M. S. BARTLETT
W. G. EMMETT
M. G. KENDALL
D. N. LAWLEY
E. S. PEARSON
E. A. PEEL
L. S. PENROSE
J. FRASER ROBERTS
A. RODGER
W. STEPHENSON
A. STUART
P. E. VERNON

VOLUME XII

Printed and Published by

TAYLOR & FRANCIS LTD.

RED LION COURT, FLEET STREET, LONDON, E.C.4

Digitized by the Internet Archive
in 2026

https://archive.org/details/british-journal-of-mathematical-statistical-psychology_1959-05_12_part-1

Vol. XII Part I The British Journal of Statistical Psychology May 1959

CLASS DIFFERENCES IN GENERAL INTELLIGENCE: I

A reply to Miss Conway

By A. H. Halsey
Department of Sociology, University of Birmingham

The hypothesis which this paper defends is that the distribution of innate intelligence among different social classes is approximately random, and that the mean differences found in tested intelligence between different social classes are due, not to genetic constitution, but solely to environment. The arguments of Conway and of Burt in favour of the alternative hypothesis are examined; and it is concluded that, with more precise figures for social mobility, their arguments would strengthen rather than weaken the hypothesis here maintained.

I. The Theory of Intelligence

In a recent paper [1] I argued the case for holding that mean class differences in measured intelligence reflect environmental factors in test performances; and in a later note [2] Mrs. Floud and I re-asserted our acceptance of the hypothesis of near-randomness in the present social distribution of innate intelligence. Miss Conway has since contributed a well-reasoned paper in support of the contrary view, namely, that class differences in measured intelligence are largely attributable to genetic causes [3]. It may therefore be useful to review our agreements and differences.

Miss Conway and I both believe in the existence of innate differences of intelligence between individuals. We both hold that these are partly reflected in test scores, and that recent educational reforms may have the effect of modifying the class distribution of innate intelligence; but Miss Conway believes that certain differentiating genetic distributions are already present. Moreover, we both distinguish between innate intelligence (I_a), intelligence as defined in the Oxford dictionary (I_b), and measured intelligence (I_c). Miss Conway thinks I took these distinctions from Burt's γ and g_a and renamed them. In fact, as I indicated ([1], p. 5, footnote 3), I took them from Vernon, who took them from Hebb.

II. Variance in Measured Intelligence

On the question of the fraction of variance in test scores which is attributable to inheritance, we differ appreciably. Miss Conway would presumably accept Burt's latest figure of 88 per cent [4]. I considered $\frac{1}{2}$ and $\frac{3}{4}$ as the limits. Here again Miss Conway imagines I had in mind a statement of Fisher's; whereas I was in fact referring to a statement of Mather's, and said so ([1], p. 25, footnote). The controversy on this question remains unresolved (cf. [5]

and [6]); and neither of our articles contributes much to it. Miss Conway brings me up to date on the evidence concerning identical twins; but, as she says, " the main results are in keeping with those already reached by Newman "—the authority on whom I relied. She does not dispose of the evidence for a correlation between environment and test score differences among identical twins reared apart. I quoted Newman's figure of 0·51 and also Anastasi's and Foley's summary of the results of correlating the I.Q.'s of foster children with assessments of the cultural level of their foster homes (0·25 to 0·52), and added my conviction that " these results underestimate the environmental effects of social class ". Miss Conway misquotes my phrase ' these results ' as ' this result ', so that it appears to refer only to the Newman quotation, and she wonders why I am so sure " apart from personal preconceptions ". I am sure because of selective placement, and the fact that the cases do not adequately represent the full range of the social and cultural scale.

Miss Conway does, however, make the important point that, *ceteris paribus*, an improvement in the environmental conditions of the relatively underprivileged should reduce the environmentally determined variance, and raise the mean scores of the ' lower ' groups relative to the ' upper ' groups. Here, as she indicates, is a fruitful line of enquiry relevant to the main issue between us; and I would underline her insistence on carefully chosen tests and samples. Burt's London results and those obtained by Floud and me in Hertfordshire and Middlesbrough are not comparable for this purpose. It is also necessary to entertain a sophisticated view of the notion of class environment. This is not, as the latter work has shown, simply a matter of crude material differences, but a complex pattern of attitudes, aspirations, etc. about which we still know far too little. It may be misleading to say, as Miss Conway does ([3], p. 177) that " most of the environmental handicaps have now been removed ".

However, the importance of all this for the present discussion is that the greater the test variance attributable to environmental factors the more likely it is that the environmental hypothesis concerning class differences which I argued is tenable. Miss Conway puzzles over the logic of my discussion of variance quite unnecessarily. I sought only to establish that the variance within social classes as compared with the variance in the total population is such as to make the environmental hypothesis possible. I did not, as she suggests, transform this argument into the conclusion that " the observed differences between social classes are more likely to be explained by environmental than by genetic factors " without the further argumentation to which we may now turn.

III. The Genetic Models

We use pretty much the same approach to the construction of genetic models of class societies, i.e. we deduce frequencies in successive generations by applying formulae of the Hardy–Weinberg type to the case of a single pair

of genes for which the heterozygous condition is intermediate. In this way we both illustrate two empirically important phenomena, viz.

(1) social classes with different means and overlapping ranges of intelligence (I_a), and

(2) the case of a small upper class with a higher mean but a lower absolute number of gifted individuals than the larger subordinate classes.

Miss Conway has no quarrel with my logic in this connection, nor I with hers except in one minor particular. She says her scheme assumes approximately equal movement in and out of her intermediate class; "otherwise as a result of the large size of this intermediate group, the average level of the population as a whole would be modified". I do not follow this. Four factors can affect the composition of a society's gene pool, selective reproduction, mutation, genetic drift, and migration into and out of the society as a whole. But since Miss Conway's model, like mine, assumes random mating within classes and equal fertility between them and ignores the other three factors, it follows that her average for the whole population must, like mine, remain constant irrespective of the size of her intermediate group or the rate of social mobility (internal genic migration).

Logic apart, Miss Conway takes serious exception to some of the sociological assumptions of my model, in particular that absolute movement between my upper and lower classes has been equal in past generations and therefore the proportion of downward movement from the upper class has been much greater than the reverse proportion. She is right that this is unrealistic. I did not take into account at this point the well known facts of differential fertility and the proportionate increase of the professional classes during the rise of industrialism. The assumptions of my model with regard to the number of classes, the type of mobility and its magnitude, the institution of marriage, etc. were quite unrealistic and the implications of my over-simplifications must be examined. But first let us note that Miss Conway's model is also unrealistic. Her three classes are no less arbitrary than my two. The manual working class comprises about three-quarters of the population whereas her lower class is less than one-third, and my upper class of 10 per cent is nearer to commonly accepted notions than her 30 per cent. Incidentally Miss Conway believes her lower class to consist "mainly of unskilled labourers" ([3], p. 177); in fact unskilled labourers constituted only 12·8 per cent of the adult male population of England and Wales in 1951, i.e. a minority of her Class III.

What matters, however, is that, as I tried to show, the refinement of my model by the introduction of more realistic assumptions leads, on balance, to a slowing down of the rate of genic differentiation between classes, and yet my over-simple assumptions strongly support the environmental hypothesis. For Miss Conway's thesis it is quite otherwise. She assumes a rate of social mobility *of the type which carries genes for high intelligence upwards and genes for low intelligence downwards* of 20 per cent both into and out of her Class I

and her Class III. This is based on Burt's studies in inter-war London ([7], [8]). Yet Ginsberg's contemporaneous study [9] yielded estimates for the current and preceding generation of a *total* social mobility (i.e. irrespective of its relation to innate intelligence) of 12 per cent and 6 per cent respectively from Class III to Class I, and of about 1 per cent in the reverse direction. Even so, Miss Conway's argument depends upon the continuance of her assumed conditions for no less than ten generations before her Class I is raised to a mean I.Q. of 110. What little is known of the history of social mobility back through the nineteenth century and beyond [10] casts very strong doubts on the possibility that these conditions obtained. So much so that, in my view, Miss Conway's discussion strengthens rather than weakens the hypothesis which she set out to demolish.

References

[1] Halsey, A. H. (1958). Genetics, social structure, and intelligence. *Brit. J. Sociol.*, IX, 15–28.

[2] Floud, J. E., and Halsey, A. H. (1958). Measured intelligence is largely an acquired characteristic. *Brit. J. Educ. Psychol.*, XXVIII, 290–291.

[3] Conway, J. (1958). The inheritance of intelligence and its social implications. *Brit. J. Statist. Psychol.*, XI, 171–190.

[4] Burt, C., and Howard, M. (1956). The multifactorial theory of inheritance and its application to intelligence. *Brit. J. Statist. Psychol.*, IX, 95–131.

[5] Maddox, H. (1957). Nature-nurture balance sheets. *Brit. J. Educ. Psychol.*, XXVII, 153–165.

[6] Burt, C. (1958). A note on the theory of intelligence. *Brit. J. Educ. Psychol.*, XXVIII, 283–289.

[7] Burt, C. *et al.* (1926). *A Study in Vocational Guidance.* London: H.M. Stationery Office.

[8] Burt, C. (1942). Ability and income. *Brit. J. Educ. Psychol.*, XIII, 83–98.

[9] Ginsberg, M. (1932). Interchange between social classes. *Studies in Sociology.* London: Methuen.

[10] Floud, J. E. (1950). Educational opportunity and social mobility. *Year Book of Education* 117–135. London : Evans.

Vol. XII Part I The British Journal of Statistical Psychology May 1959

CLASS DIFFERENCES IN GENERAL INTELLIGENCE: II

A reply to Dr. Halsey

By J. Conway
Psychology Department, University College, London

The investigations carried out by Dr. Halsey and his co-workers have revealed appreciable differences between the averages obtained by pupils from different socio-economic classes in tests of ' intelligence ' at the 11+ examination. In their view, however, the class-differences thus indicated are the result solely of differences in environmental conditions. The following paper maintains that they are largely, though not wholly, due to differences in genetic constitution. The criticisms advanced by Dr. Halsey against the genetic theory are examined in detail, and appear to arise partly from erroneous theoretical assumptions and partly from inadequate data.

I. Innate Differences between Individuals and between Classes

In Great Britain those who transformed psychology from a branch of metaphysics into a branch of natural science were members of a school of empirical philosophers, who, in their writings on educational, social, and political subjects, vigorously championed the principle of ' equal opportunity for all ', and commonly coupled this with a belief in the Benthamite doctrine that " all men are equal in mental capacities as well as in legal rights ". In the field of education there can be little question that the efforts of the Utilitarian school did much to remove the more flagrant handicaps under which the children from the poorer classes laboured; and it was the aim of the Education Act of 1944 to abolish, so far as possible, the inequalities in opportunity that still survived in the existing system of school organization.

In the highly instructive inquiry carried out by Dr. Halsey and Dr. Floud, and described in their recent book on *Social Class and Educational Opportunity* (1956), an endeavour was made to ascertain how far these efforts have been successful. The phrase ' equality of opportunity ' they define as meaning " equality of access to the secondary grammar schools ". They are, so they tell us, " aware of substantial objections, philosophical as well as educational, to our definition ". And it is clear, I think, that their own philosophical preconceptions have largely influenced both the way in which they have interpreted their results and the more detailed assumptions which Dr. Halsey has set forth as a basis of his conclusions in the preceding paper. However, in the article of mine which he now criticizes, my chief purpose was not so much to examine the philosophical aspects of their arguments as to insist that, in order to be workable in actual practice, *any* practical deductions must be compatible with the psychological facts.

The underlying problem is one on which teachers themselves have frequently expressed their views in no uncertain terms; and I venture to think more weight should be given to the impressionistic judgments which they have reached as a result of first-hand experience. It is on record that the former headmaster of a London grammar school was once asked: " Why his sixth form pupils included so many scholarship winners from poor homes and so few fee-payers from the homes of the well-to-do ? " His own explanation was that " many of the fee-payers are inherently dull or lazy; the scholarship boy is neither, otherwise he could not have passed the scholarship examination at the age of 11 ". If one may judge from the correspondence columns of the chief educational journals, that view, at any rate until recently, was typical of the vast majority of teachers, though even they are by no means unanimous. Today, now that all places in such schools are free, the problem takes a rather different form. The question most frequently asked is that which Dr. Halsey and Dr. Floud themselves have raised: " How is it that comparatively few boys from the poorer homes get to the grammar school at all ? " Nevertheless, the commonest explanation seems to be still much the same; and most teachers, I imagine, would probably reply with Professor Valentine: " Usually because they are either innately lacking in the requisite ability or else devoid of the appropriate motivation ".

However, there are still a number of educationists, and even a large number of psychologists and social writers, who, with Dr. Halsey and Dr. Floud, would vigorously challenge this view. Many of them, like Dr. Maddox, question the whole notion that mental differences are inherited or innate; or, at any rate, they would contend that, if such innate differences really exist, they are too slight to matter or too elusive to detect with any certainty. Dr. Halsey himself does not go to this extreme. As he points out (p. 1 above), both he and I are agreed on one important point, namely, on " the existence of innate differences of intelligence between *individual* children ". But he believes that there are no such innate differences in the average intellectual level of the different *social classes.* Thus the *average* boy from the lowest social category (that of ' unskilled manual workers ') is in his view quite as intelligent as the *average* boy from the highest (' professional workers ', etc.).

In the areas investigated, we are told, the proportion of boys admitted to a grammar school proved to be six times as great in the highest social category as it was in the lowest. Now, it is evident that most of the parents who are assigned to the ' professional ' class—doctors, lawyers, teachers, and the like—must have had an intelligence well above the average in order to pass the examinations needed to qualify for their several professions: they had themselves therefore been subjected to the severe process of selection; and the geneticist would maintain that this high intelligence with which the parents were endowed must have been in part transmitted genetically to their offspring, although, if it was transmitted in accordance with the multifactorial theory, he would also expect to find wide variations among the individual children. Thus, allowing

for regression, and assuming that the I.Q. of the parents in the professional class averaged about 130, we should expect the I.Q. of their children to average about 115. Similarly, if the intelligence of the ' unskilled workers ' averaged about 90, then we should expect that of their children to average about 95. These expected values agree quite well with the averages given by Dr. Halsey and Dr. Floud for children in their highest and their lowest occupational categories. According to them, however, these differences are the result solely of differences in the environmental backgrounds of the two groups: could all such handicaps be removed, then (in their view) it would be just as easy for a boy from the ' unskilled ' group to enter the university and become a senior wrangler as it is for a boy from the professional classes; and, owing to the different size of the two classes[1], senior wranglers from the former should in fact be four times as numerous as senior wranglers from the latter, not totally unheard of.

II. The Evidence from Identical Twins

How then can we decide between these two conflicting views ? Fortunately nature provides us with control groups in the shape of ' identical ' twins; and in my previous article (this *Journal*, XI, p. 186f.) I described in some detail the statistical data furnished by 42 cases in which identical twins had been reared apart from early infancy. In reply (pp. 3–4) Dr. Halsey begins by re-examining this collection of data. He considers that the figures I gave still leave " the controversy on the question unresolved ", because I do not " dispose of the evidence for a correlation between environment and test score differences ". His wording is a little vague; and there are, I fancy, really two distinct questions which are easily confused.

In his earlier paper (*Brit. J. Sociol.*, IX, p. 15f.) he dealt primarily with the question of mental versus physical inheritance. Hence it seemed to me that the appropriate answer would be given by comparing correlations for identical twins with those for non-identical twins and for ordinary siblings: these were in fact the type of correlation Dr. Halsey, Dr. Heim, and others had quoted from Newman. Accordingly in my paper I gave a detailed table of correlations, for physical measurements, mental and scholastic tests, and other assessments, obtained from (i) identical twins and (ii) ordinary twins and siblings who had been (*a*) reared in the same environment and (*b*) reared in different environments. The question that Dr. Halsey now brings to the fore is whether or not the *differences* in the test-scores can be attributed to environmental influences. And in my original report tables were also included, which were meant to furnish at least a provisional answer to this further issue. In that paper I adopted Burt's device of assessing ' economic ' (or ' social ') and ' educational ' (or ' cultural ') advantages in terms of the same type of scale as is used for assessing the intelligence and educational attainments of the children themselves, namely, a

[1] According to the figures which they quote for England and Wales just over 3 per cent of the male population aged 15 or more are in the top social group and nearly 13 per cent in the lowest.

conventional scale in which the mean is 100 and the standard deviation 15. The figures then given, related to an earlier and much smaller batch of cases. For the entire group now available the correlations between the differences in the advantages furnished by the parents or foster parents and the differences in the children's intelligence and attainments are set out in Table I.

TABLE I. CORRELATIONS OF DIFFERENCES BETWEEN TWINS REARED APART

Twins	Parents or Foster-Parents	
	Differences in material conditions	Differences in cultural conditions
Differences in		
Assessed intelligence	0·18	0·15
Group tests of intelligence	0·21	0·43
Educational attainments	0·36	0·74

Coefficients over 0·30 are significant at the 5 per cent level.

It will be seen that differences in educational attainments are highly correlated with differences in the cultural background; there is also a significant correlation—much smaller however—between the cultural differences and the differences in the scores obtained with group tests of intelligence taken just as they stand. But the correlation between cultural differences and differences in the final assessments is so slight as to be statistically non-significant with this restricted sample.

Dr. Halsey has one further criticism to make of the data for twins. " The cases," he says, " do not represent the full range of the social and cultural scale." This remark is somewhat puzzling, particularly as we are offered no definition of ' the full range '. Judged by the occupation of their parents, the pairs of twins covered quite as full a range as that covered in the surveys reported by Dr. Halsey and Dr. Floud. As explained in my article, owing to the fact that many of the cases were ascertained through personal contact, it so happened that there was an appreciable number of pairs (11 out of 42) in which one of the twins was brought up in a household which Dr. Halsey would certainly class as belonging to his highest category: the head of the household was a university or school teacher, a ' business owner ', a ' manager ', or the like. In three of these cases the child had been adopted; and here the other twin was in one case left with its mother (a charwoman, whose husband—a dock-labourer—was almost continually out of work), and in a second case sent to an orphanage: in the third case the child was left with its middle class parents, and so does not directly concern us here. In the remaining six cases, where the twins were children of parents belonging to the ' professional ' classes, one member of each pair remained with the parents. Of those who were not brought up by the parents, three were sent to an orphanage, one was boarded out with an impecunious member of the child's own family, two were boarded out with foster-mothers of the ' unskilled worker ' class, and two with foster-parents

who were of an uneducated type but in moderately comfortable circumstances, that is, they belonged to the lowest cultural, though not to the lowest economic, class: (one such pair was described in detail in the paper Dr. Halsey criticizes). With none of these eleven pairs did the difference between the assessments for intelligence exceed 6 I.Q. points; on the other hand, the differences between the assessments for educational attainments (expressed in terms of an analogous ' quotient ' scale) amounted in one case to 57 points, and in most of the others they were well over 20 points. Where verbal group tests of the Moray House type were used, the differences were usually intermediate, amounting in one instance to 18 points, but in the rest falling well below 12 points.

III. Criticisms and Earlier Inquiries

The study of identical twins enables us to compare the effects of different environments on children of the same parents and with the same genetic constitutions; the study of children brought up in foster-homes or residential institutions enables us to compare children of different parents, and therefore of different genetic constitution, brought up in the same environment. Dr. Halsey's objection to this kind of data turns on the fact of ' selective placement '. I agree that deliberate selection may seriously weaken the evidence on which so many American investigators have relied; and for precisely the same reason I refrained from quoting results obtained from foster-children in this country. But surely this objection cannot be urged against the type of British study to which I did refer, namely, those in which the children had been removed at an early age to an orphanage or residential institution. And here, if we classify the children by the occupation of the fathers, every social class was represented in almost every institution. In these cases therefore there could have been no question of selective placement.

If Dr. Halsey wishes to counter the evidence which I (as well as other investigators cited in my article) have thus adduced, one would expect him to cite analogous cases of twins and inmates of orphanages encountered during his own surveys of Hertfordshire and Middlesbrough, and show that the factual evidence in these cases contradicted my own. He does not, however, give a single instance studied at first hand. Instead he relies mainly on early American researches.

He begins by quoting "Anastasi's and Foley's summary of correlations of the I.Q.'s of foster children with the assessments of the cultural level of their foster homes ". As to these Dr. Anastasi herself observes: " the data are difficult to interpret owing to selective placement and restriction in range of foster home environments ". Of the various studies which she reviews she considers Leahy's to be the best controlled: with these the correlation of I.Q.'s with the cultural index was 0·26 for 194 foster children and 0·51 for the 194 boys and girls who were the actual children of the parents with whom they were living. Leahy's conclusion was that " heredity is more important than environment—

a conclusion similar to that reached by Burks "[1]. I fail to see how this batch of data therefore can be interpreted as supporting Dr. Halsey's view.

As regards twins Dr. Halsey still relies exclusively on early American work[2] and quotes " Newman's figure of 0·51 ". Dr. Anastasi's phrase is " *only* 0·51 " ; and, as there were no more than 20 cases in this older inquiry, a coefficient of this size is but just above the borderline for significance. Moreover, Newman himself observes that the type of tests that he used (the Stanford–Binet and the Otis group tests) would naturally be much affected by the social, and particularly by the educational, environments in which the separated twins were brought up. That, I imagine, is a point on which we should all be agreed. The really informative figures in Newman's tables are the correlations between pairs of *identical twins* reared apart and between pairs of *ordinary siblings* reared apart: with the former group the correlation is much higher than with the latter (viz., 0·67 and 0·46 respectively). Even when ordinary siblings have been reared together, the similarity of the environment does not greatly increase the correlation: it then rises only to 0·49. Thus, the effect of increasing the similarity in genetic constitution is far more marked than the effect of increasing the similarity in the environment.

IV. The Effect of Environmental Improvements

There is one more minor source of confusion that should perhaps be cleared up. So far as this country is concerned, the starting-point of the discussion, I take it, was the surveys carried out from 1922 onwards by Burt and his co-workers in the Education Department of the London County Council and the National Institute of Industrial Psychology: these early results revealed marked differences between the averages for different social classes in the I.Q.'s obtained both from adults and from children. As we have seen, Dr. Halsey contends that all such class-differences are due solely to differences

[1] Leahy's work was carried out nearly a quarter of a century ago, and is therefore open to certain criticisms which we ourselves endeavoured to avoid in our own investigations. The American studies referred to give the correlation of the I.Q.'s of the children with those of the foster-parents (0·19 to 0·24), but not with those of their own parents, though we are told we may assume that they reached the usual figure, namely 0·50 or thereabouts. The correlations of 0·19 to 0·24 may be regarded as an indication of some slight selective placement. I have taken my quotations from the latest edition of Dr. Anastasi's book (*Differential Psychology*, 1958); Dr. Halsey does not give the reference for his quotation, but it is presumably taken from the earlier edition in which Dr. J. P. Foley collaborated.

[2] In explaining why he still relies chiefly on Newman's figures, Dr. Halsey quotes me as saying that " the main results [of my own inquiries] are in keeping with those already reached by Newman ". But if he will look at the context, he will see that my comment related primarily to " researches on more general lines ", not to my own investigations, which were only mentioned in a later part of the paper; and the ' results ' referred to are the effects of environment on personality-traits as compared with its effects on other characteristics. My point was that, according to Newman, environment has least effect on physical characteristics, slightly more on ' intelligence ' (i.e. I.Q.'s measured by group tests or the Stanford–Binet scale), still more on educational achievements, and produces the greatest effect on traits of personality or character, although, as I added, " some investigators would reverse the order for the last two characteristics ", i.e. educational achievements and personality respectively.

in the environment. During the last thirty years, however, the environmental differences have greatly diminished. Both the economic and the cultural conditions prevailing in the humbler classes have undergone a vast improvement. If then environment was the chief source of the I.Q. differences between one class and another, we should naturally expect that the I.Q. differences would likewise have diminished. Partly to show that there had been practically no change I reprinted in the first of my tables " the average I.Q.'s for different occupational classes " as given by Burt for London 30 years ago and compared them with the corresponding figures given by Dr. Halsey and Dr. Floud in their recent surveys. In both the areas which they investigated their figures were surprisingly similar to those which Burt had obtained. There was certainly no sign of any appreciable rise: indeed, their figure for the working classes was actually a fraction lower.

Dr. Halsey replies that " Burt's London results and those obtained by Dr. Floud and [himself] are not comparable for this purpose ". Were that so, the similarity would be still more astonishing. I should agree that the data might not be regarded as strictly comparable were we comparing the performances of individual children. But, when we compare the groups, the minor differences appear to cancel out; and, in order to show that the results could fairly be compared, I pointed out that Moray House tests (similar to those used by Dr. Halsey and Dr. Floud) had been applied to a sample of the children tested by Burt and myself in 1950. Had my sole object been to compare averages for the pre-war and post-war periods, I could, of course, have given averages for London children obtained in a recent inquiry in 1950 where precisely the same tests were used on both occasions. (The figures were reported in a paper by Burt, presented last year at a meeting of the British Psychological Society: I reproduce them for comparison in the table below, copied from tables

TABLE II. AVERAGE I.Q.'S FOR CHILDREN FROM DIFFERENT OCCUPATIONAL GROUPS

Class	(A) London 1922–27	(B) London 1950	(C) S.W. Herts 1952	(D) Middlesbrough 1953
1. Highest Administrative	120·3	122·1	—	—
2. Professional	114·6	115·4	113·0	115·2
3. Clerical	109·7	109·0	109·2	108·2
4. Commercial	104·8	102·2	103·7	104·4
5. Skilled	101·5	98·3	100·1	99·0
6. Unskilled	95·1	95·8	97·1	96·5
Average for working classes	98·3	97·1	98·2	

NOTE. In 1925, 8·4 per cent of the London elementary school population were living in 'extreme poverty' (according to Rowntree's definition) and 22·3 per cent in 'poverty', most of the cases being found among the 'working classes' (groups 5 and 6); in 1950 the corresponding figures were 2·3 per cent and 11·4 per cent. In Hertfordshire (1952) " material prosperity in all classes was high enough to prevent its influencing results obtained ". In Middlesbrough (1953) " even the majority of the unskilled received wages of £7 10*s.* or over ".

circulated at the meeting.) However, it seemed that my argument would carry far more weight if I gave the figure Dr. Halsey himself had printed in his own paper, particularly as my main point was to demonstrate that the class-differences which *he* had obtained were quite as large as those found in the original London surveys.

It is plainly impossible to accept any longer the older arguments put forward by the environmentalists, namely, that, since the class-differences in I.Q. are in exact correlation with the class-differences in financial status, therefore the real causes of the low I.Q.'s found among the ill paid classes must be poverty and its various concomitants. Dr. Halsey, while still adhering to an environmentalist theory, explains that " it is necessary to entertain a sophisticated view of the notion of class environment ". The causes to be invoked therefore are not simply " a matter of crude material differences ", but other more elusive differences " about which we know far too little ". Yet if so little is known about them, how can we be so sure that they are in fact responsible for these wide differences in ' intelligence ' ? I cannot help feeling that the impartial reader will be tempted to say that it is the reasoning rather than the ' view ' that is really ' sophisticated '.

V. The Joint Effects of Individual Variation and Social Mobility

In the third section of his paper Dr. Halsey very generously accepts my criticisms of the figures he had adopted for the interchange of intelligent and unintelligent members from the different social classes; and he admits that it will now be necessary to examine the effects of the ' over-simplifications ' involved in his previous model. It will be certainly instructive to see what conclusions he himself arrives at after he has completed this examination. At the same time, however, he criticizes my own model on the ground that it is ' unrealistic '. Here I am not clear whether his objections relate to the algebraic or the numerical assumptions involved. As regards the algebra, that is merely an adaptation of Burt's earlier model; and I shall therefore leave the defence of this aspect of the question to him. As regards the numerical values that I adopted the obvious reply is this. It is the very essence of a model to be schematic, and to that extent ' unrealistic '. My own ' model ' therefore was in no way intended to supply a realistic description of the actual way in which different social classes had developed either in this or any other country: that is a question for the historian. My object was merely to start from perfectly general premisses based on what is known of human genetics and of social mobility, and to discover by a purely mathematical argument whether the ultimate result would be a social structure in which all the classes would have much the same genetic composition or whether it would reveal a trend towards class-differentiation. The insertion of arithmetical values instead of symbolic values was adopted merely to render the argument a little more intelligible; and, provided the figures chosen conform to the initial postulates, could entail (at least so far as the main issue is concerned) no real loss of generality. For

the same reason simplified proportions were adopted for each of the three classes and rounded averages for the rate of interchange. However, to avoid these issues, let us keep to Dr. Halsey's own model and disregard mine.

If I understand him rightly, he believes that the main effect of " introducing more realistic assumptions " will be a change in the " rate of genetic differentiation between classes ", probably ' on balance ' a change to a slower rate. With that I should be disposed to agree. Yet I cannot see how this of itself could yield results that " strongly support the environmental hypothesis ". What are Dr. Halsey's own conclusions ? Starting with three main classes in which the mean I.Q.'s are all 100, we find that, after about seven generations, low intelligence is bred out of the highest class altogether, while the mean I.Q. for the lower class decreases to 98. That leaves a class difference, due to genetic constitution, of as much as 22 I.Q. points. If then the actual rate had been twice as slow, the *final* outcome would not have altered: it would have merely taken 14 generations instead of 7 to reach it, say about 500 years instead of 250.

However, as I observed in my previous article, it is difficult to see what precisely is the justification for the particular model that Dr. Halsey wishes to substitute. I supposed that his basic assumptions were taken from Fisher's discussion of the problem. But in this, it now seems, I was mistaken. Nevertheless, whether we recognize it or not, it is Fisher's methods and formulae that we are all of us using; and it is to Fisher that we owe the very notion of dealing with the problem by means of a mathematical model based on Mendelian principles.

The essentials of Fisher's techniques are contained in his classical volume on *The Genetical Theory of Natural Selection* (1930). The last eight chapters bear very closely on the issues which Dr. Halsey has raised. As Fisher observes, his " deductions respecting Man are strictly inseparable from the more general chapters " (i.e. the mathematical portions of the book). Nevertheless, as the nature of his arguments and conclusions does not seem to be as well known as it should be, it may perhaps be worth while summarizing them here, using, so far as I can, his own words.

He begins by pointing out that " among the problems presented by the social evolution of Man the most conspicuous is that of the decay and ruin of all civilizations previous to our own ". He then examines the various hypotheses put forward to account for this unexpected phenomenon; and he argues that, since the knowledge, skill, and power which are acquired by men in civilized societies do not suffice to defend them against peoples who are relatively uncivilized, the real cause must lie in a change in their innate qualities, partly physical perhaps, but mainly psychological. As all such innate qualities are transmitted in accordance with Mendelian principles, they must exhibit wide variations from one individual to another. Now, during the greater part of their history the policy adopted by each nation is the policy of its effective rulers; and these rulers, whether they have emerged within the nation or whether they

have imposed their will as alien conquerors invading the territory from without, must have achieved their position largely[1] because they were themselves innately endowed with superior physical, mental, or moral qualities. It follows that " the immediate cause of decay must have been the degeneration or the depletion of the ruling class ".

" Biologically," he says, " social classes are of importance by controlling mate-selection, through the influence of prevailing opinion and the opportunities for social intercourse which they permit. Social classes thus become genetically differentiated, like local varieties of species, the differentiation being determined primarily by social promotion or demotion." Among primitive societies " the importance attached to kinship " has been of special efficacy by giving a selective advantage to qualities recognized as socially valuable; and as a result " the higher qualities of man, and his appreciation of them, seem ascribable to social selection in this type of society ". Indeed in many instances the importance of avoiding intermarriage between different social classes or different ' races ' (so-called) has been felt so strongly that this kind of mating has been inhibited by drastic legal penalties. Where any mingling does take place, then (according to Mendelian principles) the *average* quality of the mixture will be intermediate, but the *variability* will be far greater than that of either of the original groups. This increase in variability furnishes " great possibilities for the action of selection .... If the selection is beneficent, the result will be superior to both the previous groups; if unfavourable, it will be far more rapidly disastrous." As time goes on the problem becomes more complicated, because, as Galton originally pointed out, there is a tendency for individuals, classes, and nations to reproduce their kind more slowly when they become more prosperous. Thus in various ways the members of civilized societies, without quite realizing what they are doing, tend to modify the course of their own evolution. And Fisher's final conclusion is that " the genetic composition of existing populations, graded both in social ability and in effective infertility, presents grave difficulties, which only a people capable of a deliberate and intentional policy can hope to overcome ".

Similar views have been put forward still more recently by Prof. Darlington; and, as I stated in my paper, both the general assumptions and the general procedure that I myself adopted seemed quite in line with theirs. My chief purpose was simply to apply them to a particular set of empirical data. If therefore Dr. Halsey wishes to reject the inferences that follow, he should, besides criticizing my own methods and assumptions, indicate where in his view the arguments of these eminent authorities are at fault.

[1] As history amply indicates, the possession of superior numbers or of superior military equipment is of itself by no means decisive.

Vol. XII Part I The British Journal of Statistical Psychology May 1959

CLASS DIFFERENCES IN GENERAL INTELLIGENCE: III

By Cyril Burt
University College London

The argument put forward in the following paper is based on three main premises, for which ample evidence is now available: (i) individual differences in 'general intelligence' are transmitted in accordance with the multifactorial theory of inheritance; (ii) in our own country for thirty generations or more, there has been a discernible tendency for individuals endowed with high intelligence to rise to a higher social class and for those of low intelligence to drop to a lower; (iii) with minor fluctuations, this twofold movement has increased at an accelerated rate from almost negligible amounts to the comparatively large proportions reported in recent sociological surveys. From these three propositions it is concluded that there must now be appreciable differences in the genetic composition of different social classes, and that these differences will show themselves most plainly in differences between the mean level of intelligence obtaining within each class. But it also follows that, within each class, there will be wide individual variations, so that each class remains highly heterogeneous; and these differences in innate ability will be still further magnified by environmental differences and by innate differences in temperament.

The conclusions thus deduced appear to be fully confirmed by surveys carried out with tests of intelligence and other devices. But the test-results themselves by no means form the only ground, as is so often supposed, for accepting the conclusions thus reached. The various objections urged against the foregoing arguments are discussed; and their practical consequences briefly indicated.

I. Alternative Hypotheses

The Genetic Hypothesis. Dr. Halsey in his recent papers [13, 16] has raised afresh the problem of class-differences; and even those who do not agree with the conclusions that he and his collaborators have reached will be grateful to him for stating in explicit detail the case for a purely environmentalist theory. That theory has been tacitly accepted by a number of educational and social writers who have discussed the current scheme of educational organization in this country, particularly in connection with the 11 plus examination. But the reasons commonly advanced appear to be based rather on the social ideals or the subjective preferences of the critics than on any first-hand examination of the evidence supporting the opposite view.

Deductive Arguments. The evidence itself may conveniently be treated as consisting of two main parts—deductive or *a priori* and inductive or *a posteriori*. What may be called the deductive argument is based primarily on the theory of mental inheritance which I believe practically all geneticists would today accept—namely, that individual differences in graded characteristics, such as

intelligence, are transmitted mainly in accordance with the principles of polygenic or ' multifactorial ' inheritance. Now all organisms, whether plants or animals, exhibit special devices which tend either to limit or to favour certain types of breeding-system. In man the two restrictions on mating which characterize nearly all cross-breeding organisms are, as a rule, rigorously observed—namely, those that discourage the more extreme forms of in-breeding and of out-breeding respectively. But in most human communities automatic control of breeding is replaced by social and intellectual control[1]; and within certain fairly wide but variable limits human groups often swing over from in-breeding to out-breeding and back again in ways that are traceable in no other organism. The result is the formation of groups within groups, and in particular of what are popularly called ' races ' and ' classes '. Among the more civilized communities there have thus arisen, within the same broad area, interfertile classes, which, under a common social organization, work together, but do not as a rule breed together; and this type of social structure has been largely responsible for the rapid advancement of the societies in which it is found. In those communities which have survived for an appreciable time, however, some degree of interchange has nearly always prevailed: individuals in the lowlier classes, who have been exceptionally endowed with the qualities of physique, intelligence, or character distinctive of the higher classes, have succeeded in crossing the barriers and rising above their initial station, while individuals in the higher classes who have been conspicuously lacking in such qualities have tended to drift to a lower level.

In our own country a clear distinction between different classes is discernible well before the Anglo-Saxon and the Norman conquests; and from about the twelfth century onwards there is ample record of a growing tendency for persons born in one class to rise or fall into another as a result of their special merits or deficiencies. To begin with, no doubt, such interchanges must have been comparatively rare; and their number has fluctuated greatly according to the customs and conditions of the age in which they lived: but on the whole it has steadily increased up to the present day. And fairly precise information is now available about the actual extent of such movements during the last two or three generations.

However, as Fisher's investigations have clearly shown, such tendencies and their effects can only be discussed intelligently if they are expressed in mathematical form. Accordingly, on the basis of the evidence now available

[1] The first to emphasize the importance of these two processes were the early anthropologists, notably J. F. McLennan, to whom we owe the terms ' endogamy ' and ' exogamy ' (cf. *Primitive Marriage*, 1865, *Studies in Ancient History*, 1876). Bateson appears to have been the first to recognize that these two types of breeding in man are merely special cases of the two wider principles traceable in the breeding systems of plants and animals; but Darwin's studies of sexual selection had already indicated the importance of mating-systems generally in the evolution both of man and of the humbler organisms. It is unfortunate that the various studies of mating-systems undertaken by geneticists on the one hand and anthropologists and sociologists on the other have been carried out with so little reference to the common implications of each other's work.

regarding the transmission of genetic differences and the nature of the interchanges between the different classes, I ventured to formulate in quantitative terms a hypothetical model or scheme, from which inferences might be drawn about the genetic constitution of different classes at any particular stage.

Empirical Evidence. The final step in the argument—what I have called the *a posteriori* or inductive stage—is to test the corollaries so deduced from the model by comparing them with the actual results obtained from surveys carried out by means of mental tests and similar devices among representative samples of the adult and the child population. Several such surveys were undertaken in London and elsewhere from 1914 onwards. As reported in my various memoranda [2, 5] wide differences were found between the averages of the individuals from different socio-economic groups. The children were subdivided into seven main classes: class I consisting of children of parents in the higher professional or administrative occupations, class VII of unskilled casual workers. And the average I.Q.'s ranged from 120 for class I down to 89 for class VII [4]. Similar surveys were carried out on a more limited scale with adults by my colleagues and myself at the National Institute of Industrial Psychology in connection with inquiries into the possibility of vocational guidance. These revealed analogous and still wider differences between the averages for the various socio-economic groups ([3], p. 16). The correlation between the children's I.Q.'s and the socio-economic status of the parent was 0·32[1]. And it was argued that well over 50 per cent, and in all probability as much as 75 per cent, of the individual variance was attributable to inheritable differences[2]. Thus the empirical results seemed entirely consistent with the deductions drawn from the theoretical premises. The data, however, on which I myself relied in these earlier articles and memoranda were mainly gathered during my work in London schools before the war. Miss Conway [9] has endeavoured to supplement them with more recent material and to simplify the mathematical formulations. Once again the general outcome of the analysis seems fully to confirm the genetic hypothesis.

[1] This is appreciably higher than the figure obtained by Gray and Moshinsky (Hogben, [17], p. 416) in the later London inquiry, quoted by Halsey and Floud, viz. 0·25; that inquiry, however, was based on tests originally constructed for American children. The figure obtained by Duff and Thomson [11] in Northumberland was slightly higher (0·28). But there the marks obtained from the group tests were taken just as they stand without corrections for incidental disturbances inevitable with young examinees of only 10 or 11: this must have somewhat reduced the size of the correlation so computed.

It was these early surveys with intelligence tests that first clearly demonstrated the vast amount of undetected talent lying hid among the children of the lowlier classes. In 1921 it was found that over 60 per cent of those who are capable of a university education never obtain it. This and similar findings seem to be entirely overlooked by those who criticize such tests on the ground that they discriminate *against* children from the ' unprivileged ' groups.

[2] In their reply to Dr. Warburton [13] Dr. Halsey and Dr. Floud take him severely to task for supposing that " intelligence tests measure *pure innate ability* ", and add that " his faith in twin studies for this purpose cannot be taken at its face value ". But neither Dr. Warburton nor those who have studied twins ever supposed that intelligence tests measure ' pure ' innate ability: Dr. Warburton merely claimed that what is measured is ' largely ' innate [21].

The Environmental Hypothesis. Nevertheless, this genetic hypothesis and the arguments on which it is based have been vigorously contested by Dr. Halsey and his co-workers. So far as our present problem is concerned, their own conclusion is most succinctly summed up in their reply to Dr. Warburton's criticisms in the last issue of the *British Journal of Educational Psychology* [21, p. 290]. They cannot, they say, " accept the inference that differences in the mean intelligence test scores of children from the various social classes reflect the social distribution of intelligence A or innate ability; on the contrary, we accept as most plausible the hypothesis of near-randomness "—i.e. of the virtual equality of the true means for the different classes. And if I interpret Dr. Halsey's present paper correctly (pp. 1f.), his criticisms of our conclusions turn upon three main points: (i) the evidence for class-differences in intelligence as based on test scores; (ii) the genetic model assumed in accounting for these differences; and (iii) the estimates for social mobility.

II. The Evidence from Mental tests

The Analysis of Variance. Dr. Halsey apparently assumes that the evidence for class-differences in innate intelligence is based solely, or at any rate chiefly on an examination of test-results like those which he and his co-workers have analysed in their book *Social Class and Opportunity*. Their data were collected during the course of a series of instructive inquiries into the actual working of current methods of selection in Middlesbrough and South-West Hertfordshire. Now, neither he nor his colleagues wish to deny the existence of a genetic component in individual differences. What they do deny, at any rate in their book, is that the scores obtained from intelligence tests such as those that they describe can in any way represent these genetic differences: " intelligence tests ", they say, " reflect current education organization and practice; and measured intelligence represents a set of *acquired* aptitudes " ([12], p. xix, my italics).

Genetic and Environmental Differences. In his present paper, however, Dr. Halsey himself seems to retreat a little from this rather extreme position. He begins by emphasizing that he " believes in the existence of innate differences of intelligence between individuals ", and goes on to admit, with Miss Conway, that " these are partly reflected in test scores ". What he now criticizes, it appears, is the way in which I analysed the total amount of variance found with such tests. After removing a component attributable to ' unreliability ', I endeavoured to analyse the variance into that which was attributable to variations in the class means and that which was attributable to the variations of the individual members about the mean for their particular class; and I introduced a cross-classification dividing the variance in each case into a part attributable to differences in genetic constitution and a part attributable to environmental influences. The main differences between Dr. Halsey and his colleagues on the one side and Miss Conway and myself on the other arises, so he tells us, over " the proportion of the variance in the test-scores which is attributable to inheritance ". He considers " $\frac{1}{2}$ and $\frac{3}{4}$ as the limits "; but he supposes that Miss

Conway and I would adopt a " figure of 88 per cent ". But here there is an obvious misunderstanding. In the paper he cites, what I actually said was that " with *adjusted* assessments for intelligence, obtained in the way we have described, about 87 per cent would be attributed to genetic constitution ", but " with the scores for group tests taken as they stand *only about* 77 *per cent* ". Thus for our own group tests my figure does not differ appreciably from the higher of the two figures that Dr. Halsey himself suggests.

The Nature of the Tests. The tests referred to in this quotation, however, included both verbal and non-verbal tests, and were designed and standardized specifically for London children. But the tests Dr. Halsey has in mind were intended, not for the purposes of research, much less to study the effects of intelligence as distinct from those of environment, but as part of an 11 plus examination, where the object was to estimate the suitability of the candidates at the time of the examination for education at a grammar school; and they were standardized, not in the areas in which they were used, but by a team of psychologists at Edinburgh. Indeed, to avoid any such misunderstanding, Moray House expressly designates them, not ' tests of intelligence ', but ' tests of verbal reasoning '. For scores obtained from tests such as these I should be well content to accept Dr. Halsey's assessment of the ' variance attributable to inheritance ', namely, 50 to 75 per cent. Those are in fact the limits I myself suggested in discussing tests of this type; and they agree with the proportions given by Thomson, who was responsible for the general plan of the Moray House tests. His argument was that " they would presumably have a correlation of between 0·70 and 0·85 with innate ability, and thus (taking the squares) between 0·50 and 0·75 of the total variance would be attributable to innate differences ".

Class Differences and Individual Differences. The real difference between Dr. Halsey and myself therefore would seem to be this. We both believe in a moderately high correlation between *individual* measurements and the underlying innate differences between the individuals; but, whereas I should maintain that this implies a consequential correlation between the *class* means and certain innate differences between the classes, Dr. Halsey denies the inference thus drawn. His argument apparently runs as follows. An appreciable portion of the variation in the *individual* measurements is due to environmental and not to innate differences; it is therefore conceivable that the variation in the *class* measurements may be due *entirely* to the environmental influences: and this, as he now says, may be " transformed into the conclusion that the observed differences between social classes are more likely to be explained by environmental factors " by means of ' the further argumentation ' set out in section III of his present paper.

Before we turn to this ' further argumentation ' let us consider what precisely it will have to prove in order that this additional conclusion shall follow. In my original paper the several occupational classes were arranged in order from I to VII according to " the apparent difficulty of the mental processes required either to carry out the occupation efficiently or to gain

entrance into that occupation"; and with this order of 'occupational difficulty' the order given by the average class measurements for the children's intelligence showed a perfect correlation. The same is true of the average class measurements given by Dr. Halsey and Dr. Floud for each of the areas they investigated. Furthermore, both Dr. Halsey and I agree that there is a correlation of at least 0·70 between the individual measurements of the intelligence of the children and their innate ability. Now, if the children's ability is innate, it must also be correlated with the innate ability of the parents; and for this, I take it, we may assume that Dr. Halsey would accept the usual coefficient of 0·50. If so, the only point over which we can differ is whether or not there is also a correlation between the innate ability of the parents and the difficulty of the occupation in which they are engaged. I myself should argue that there is ample evidence to prove that this correlation is positive and moderately high. But if Dr. Halsey's additional conclusion is to hold good, his 'further argumentation' must demonstrate that *this correlation is zero.*

For example, in Middlesbrough the class mean for children of professional parents is nearly 20 I.Q. points higher than the class mean for children whose fathers are unskilled labourers. But this, in Dr. Halsey's view, is due solely to the superior intellectual environment of the children belonging to the professional group; so far as *innate* ability is concerned, the mean for each class—the professional, for example—is (on his hypothesis) equal to that of any other class—the unskilled labourer. Indeed, according to Dr. Halsey and Dr. Floud, the very fact that there is a positive correlation between the average for the tests and the occupational class of the children proves "the influence of environment on intelligence test scores" [12, p. 143].

But by itself, of course, a positive correlation between x and y cannot possibly *prove* that x causes y or that y causes x. We need, besides the mere statistical figure, some *a priori* theory as to the causal interrelations between the two correlated variables. That I endeavoured to supply in my own discussion of the problem by what I have called the *a priori* arguments. As we have seen, these arguments consisted of a schematic deduction, expressed in mathematical form, showing the consequences of the two main causative factors at work—first, the multifactorial transmission of genetic differences, and secondly, the repeated transference of individuals endowed with an intelligence either appreciably higher or appreciably lower than the average intelligence of their own occupational class, into another class for which they are better suited. Dr. Halsey has proposed a rather different model, leading, he believes, to the opposite conclusion. Our next task therefore is to examine a little more closely the factual premises on which our models were based and the validity of the deductions that follow.

III. The Genetic Models

The Effect of the Simplifications. Dr. Halsey begins by pointing out that the model described by Miss Conway is highly simplified: it ignores not only such factors as mutation and genetic drift, about which little or nothing is known,

but also the facts of assortative mating and of migration into and out of the community as a whole, about which far more evidence is available. As Miss Conway herself observes, however, my own model made some attempt to allow for the last two factors as well as for differential fertility; yet the results simply showed that, even when the allowances err on the generous side, these factors make little difference to the main conclusion. The chief effect of assortative mating is to speed up the process. Marked differences in class-fertility appear to be a relatively recent phenomenon; and so far there is little or no evidence that their effects have as yet greatly altered the level or the distribution of intelligence. In any case, Dr. Halsey's model included the same simplifications.

The Subdivision into Classes. There is, however, one further simplification over which we disagree. For the purpose of arithmetical computation, it will be remembered, both Miss Conway and I adopted the familiar working model, based on a single pair of allelomorphic factors, *A* and *a*, which, with random mating, would by Mendel's laws produce three types, *AA*, *Aa*, and *aa*, in the proportions of 1 : 2 : 1 (cf. [19], p. 173). Corresponding to these three types, we assumed a triple division of the whole community into three main classes. Once again Dr. Halsey has adopted much the same mode of simplifying the arguments; but he strongly objects to the way we have subdivided the total population.

Now this method of simplifying the deductions derived from the theory of multifactorial inheritance is in no way a mere arbitrary device, as he appears to assume. It has in fact been demonstrated mathematically that, if a large number of genes determine the variations in a graded trait according to principles of multifactorial transmission, and if the variations in question are then grouped to form three main types in the way described, the frequencies obtained will in fact be the same as those deduced for a qualitative trait depending on a single pair of genes of the kind specified. Since the frequency distribution for intelligence is approximately symmetrical, we are justified in adopting a symmetrical distribution for the constituent types. Assuming that the initial frequencies are 25, 50, and 25 per cent respectively and the S.D. 15, the average I.Q.'s will be 80, 100 and 120. Dr. Halsey takes the same I.Q. values for the means of his three classes; but he makes his lower class 9 times as large as the upper. This seems a needless departure from the factual data; and destroys the strict validity of his mathematical simplification of the multifactorial theory.

Dr. Halsey explains that his frequencies for the three main classes are " nearer to commonly accepted notions " than ours, since " the manual working class comprises about three-quarters of the population ", and the phrase ' upper class ' popularly refers to the ' upper ten ' (i.e. upper 10 per cent). But we were here concerned, not with a classification based on occupations or on social prestige, much less on " commonly accepted notions " regarding social groups, but solely with the formulation of a genetic model. On the other hand, Dr. Halsey's threefold classification is not in keeping with the theoretical requirements of his genetic assumptions.

IV. Social Mobility

The Rates of Interchange. However, the most important difference between us arises, as he says, out of our divergent assumptions in regard to the rates of social mobility. Dr. Halsey's model assumes that " the proportion of downward movement has been much greater than the reverse proportion ". On the other hand, my final model, like Miss Conway's, assumed that in the long run the upward movements and downward movements would be about equally balanced. The corollaries of other theoretical schemes were also considered; and it was not difficult to show that any marked and sustained inequality in either one direction or the other would produce results quite inconsistent with the observable facts.

Miss Conway's figures, as Dr. Halsey remarks, were " based on Burt's studies in inter-war London ". In his view, however, these early figures greatly exaggerated the amount of social interchange. " Ginsberg's contemporaneous study", he says, " yields estimates of *total* mobility of 12 per cent and 6 per cent respectively from class III to class I, and of about 1 per cent in the opposite direction." But these estimates would mean that, during the period in question, the movement from the *lower* class was " much greater than the reverse proportion". I therefore fail to see how they would support Dr. Halsey's assumption. At the same time they are not necessarily incompatible with our own assumptions. Our figure of 20 per cent, unlike the figures of Ginsberg quoted by Dr. Halsey, was intended to include the amount of movement not merely from the lowest class (III) direct into the highest (I), but from class III *and class II* into class I (and vice versa). Ginsberg himself remarks that " the movement from class II to class I is naturally greater " than the movement from class III to class I. However, like most other investigators of social mobility, he was much more interested in studying how far deserving children from the lower social ranks succeed in rising to the higher social ranks than in studying how far the undeserving drop from the higher ranks to the lower; and, if Dr. Halsey holds that Ginsberg's data underestimate the downward movement, I should be tempted to agree.

The Present Rate of Social Interchange. But by far the most reliable information on the subject is now to be found in the papers published by Glass and Hall. Let us therefore examine the figures given in one of their latest and most comprehensive reviews ('A Study of Intergeneration Changes in Status' [14], Table 6, p. 186). They started with a sevenfold classification, and eventually reduced it to five classes. For our purpose it will be convenient to pool them still further, so that they may to some extent correspond with our threefold genetic subdivision. For the period 1910 to 1930 the relative sizes of the three classes, as given in their tables, were as follows: class I, 25·2 per cent; class II, 47·6 per cent; class III, 27·2 per cent. In the last decade the slight asymmetry was reversed, and the proportion in class II had risen to 53·2 per cent. Thus for the period covered by my own investigations (1915 to 1930) their proportions are very much the same as mine, namely 25, 50, and 25 per cent. According to the grading proposed by Glass and Hall, class III

includes all "semi-skilled and unskilled manual workers", class II "skilled manual workers and those carrying out routine types of non-manual work", and class I not only "professional, administrative, executive, managerial, and supervisory workers," but also practically all forms of non-manual work that are not of a routine type.

I have averaged their figures for the different dates, using the figures on p. 93 as weights, and obtain the following distribution (Table I). The figures

TABLE I. DISTRIBUTION OF SONS ACCORDING TO THEIR OWN AND THEIR FATHER'S SOCIAL CLASS

Status of Fathers Class	Present Status of Sons Class			
	I	II	III	Total
I	51·7	34·5	13·8	100·0
II	23·3	46·9	29·8	100·0
III	13·7	36·9	49·4	100·0

are to be interpreted as follows: of sons who had fathers in class I, 51·7 per cent—roughly one-half—remained in the class in which they were born, while the rest moved down to a lower class; similarly, of those who were born in class III just over one-half (13·7 + 36·9 = 50·6 per cent) moved into a higher class.

Thus, for the period Dr. Halsey has in mind, the rounded figures we assumed so far from exceeding those obtained in actual inquiries, are in fact much lower. The reasons for adopting these lower proportions were as follows. We were concerned, not with the actual movements from one class to another, much less with the actual movements in the twentieth century, but simply with a schematic model representing the general trend of social interchange from the early dawn of civilization in these islands. For this purpose it would plainly be wrong to assume that the interchange has been going on at the same high rate from the very start. My own assumption was that it had developed more or less progressively from a primitive feudal or even caste-like condition, in which there was practically no interchange to conditions obtaining at the present day. Moreover, not all the movement could be attributed to the qualities in which we were interested. Roughly speaking it may be said that the formation of an élite or upper class is determined, in the course of a nation's evolution, chiefly by physical force at the outset[1], then by blood-relationship, and later by

[1] I am here confining myself to the historical stage alone. It would, I imagine, now be generally agreed that the class-system is much older than the nation or state. The most primitive communities of which we have knowledge—the so-called 'foodgatherers'—had no class-system. A class-system only appears among the so-called 'food-producers', and seems to have arisen out of the need for a controlled organization of labour as a pre-requisite to the successful development of a large agricultural community. In Sumer, Egypt, India, Mexico, and Peru the emergence of a ruling group is associated with the emergence of a routine and a ceremonial governing the provision of an adequate food supply, and particularly with the invention of an efficient calendar, which in the words of an eminent anthropologist must have been "one of the greatest intellectual

property or wealth, and finally by mental efficiency. Most of the time no doubt each of these factors are operative simultaneously though in varying proportions, and at every stage mental efficiency must have played *some* part; but its importance must have steadily increased until at the present time it now preponderates over the others.

The Average Rate of Social Interchange. Had the rate of increase been uniform from the very beginning, the average amount of interchange would have been about half of the final figure. But, as Miss Conway has indicated, we assumed that the rate of increase itself increased at an accelerated pace, though, of course, with a continual ebb and flow. Judging by the figures available, the total effect appeared to be equivalent to an average mobility of about 20 per cent in and out of class I, and approximately the same percentage in and out of class III. In order to abridge and simplify the exposition, Miss Conway kept to the same figure throughout, since her aim was merely to illustrate the basic principles. But her main conclusions are precisely the same as mine: namely, that the average I.Q. of the upper class (defined as above) would gradually increase, but would never reach an I.Q. of 120, because it would never consist wholly of the type *AA*; and similarly that the average I.Q. of the lower class would gradually diminish, but would never reach an I.Q. of 80, because it would never consist wholly of the type *aa*.

Dr. Halsey's final objection is that, under the conditions assumed, the time required to achieve the present condition would have been incredibly long. With Miss Conway's model, as he observes, it would have taken " no less than ten generations before her class I was raised to a mean I.Q. of 110 ". With my model it would have taken still longer. Dr. Halsey thinks that " what little is known of the history of social mobility casts very strong doubts on the possibility that these conditions obtained ". Yet I fail to see how he comes to this reading of history[1]. After all, ten generations would take us back only

feats ever accomplished ". Thus the ruling group at the outset was a highly intelligent group. Later, when they migrated to other territories, they imposed themselves as conquerors; and thus formed a military rather than an intellectual aristocracy. If this hypothesis is correct, then the influence of innate intelligence on the formation of classes must have started well before the dawn of history: (cf. [1] and refs.).

[1] In their reply to Dr. Warburton [13], Dr. Halsey and Dr. Floud again place the historical argument in the forefront. " We take the view," they say, " that the polygenetic nature of intelligence and the history of social mobility in this country exclude the possibility that the present distribution of ' intelligence A' (i.e. innate intelligence) between social classes is such as to give mean differences between them of the order given by studies of measured intelligence." But they give no indication of the historical facts they have in mind, and they cite no references to explain their view.

From the remarks in the opening chapters of their book and from one or two phrases in Dr. Halsey's present paper it would almost seem as if they assume that mental efficiency had played no appreciable part in the interchange between different classes until the last half-dozen generations. That perhaps might be the impression one would gain if one merely studied the history of social and educational progress " through the nineteenth century " and just beyond. But first-hand study of earlier records will quickly show this popular notion to be erroneous. Thus at the older universities, down to about the close of the 16th century, we find that the majority of scholars are

to the end of the 17th century; and there was considerable social mobility long before that. Toward the close of the feudal period the founding of the universities and the grammar schools, the growing importance of the legal and clerical professions, and above all the rapid extension of trade and the consequent emergence of the middle class, indicate a steady increase in the influence of sheer mental efficiency in determining a man's ultimate position in the social hierarchy (cf. [7, 15]). Legendary success-stories, like that of Dick Whittington, imply that the notion of upward mobility was already familiar in the 14th century[1]; and moralists, long before Hogarth, emphasized the ease of mobility in the opposite direction (see [20] and refs.).

But in my opinion the actual history of the subject has only an indirect connection with the main issue over which Dr. Halsey and Dr. Floud differ so widely from Miss Conway and myself. Miss Conway and I were not attempting to reconstruct the history of social interchange or to deduce its actual speed. Our problem was a problem in psychogenetics; and for this purpose it would be sufficient if we could safely assume (i) that for many centuries—probably nine or more—there was at least some degree of interchange between the various social grades or ranks, (ii) that this interchange was determined largely by the presence or the lack of certain innate qualities making for social efficiency, and finally (iii) that it rose towards the close of the 19th century to the high figures given by Glass and Hall. The actual proportions, and the actual rates at different points of historical time, seem to us to be entirely irrelevant.

The Influence of Character Qualities. Part of the argument brought forward by Dr. Halsey and Dr. Floud turns upon the fact that among the children in the lower social classes quite an appreciable proportion of those who, according to our hypothesis, should be innately endowed with high intelligence, nevertheless fail to get into the grammar school, or, when they do, fail to reach a university. The underlying assumption apparently is that the only other

matriculated as '*pleb. fil.*', an increasing number as '*cler. fil.*' (son of a man in holy orders) and very few as '*arm. fil.*' (son of a man entitled to armorial bearings): *gen. fil.* (of 'gentle' birth) is rare at first, but forms the bulk of the entries from the middle of the 18th century. As a result during the 14th to the 17th centuries it was much easier for a poor boy of exceptional intelligence and character to work his way up to the highest levels than it was in the middle of the 19th century. To rise from the ranks of the intermediate classes was easier still. Nicholas Breakspear (Pope Adrian IV) was the son of a shepherd, Chaucer of a vintner, Wolsey of a grazier, Thomas Cromwell of a blacksmith, Knox of an East Lothian peasant, Spenser of a clothier, Marlowe of a shoemaker, and Pepys of a tailor. Havelock Ellis in his *Study of British Genius* comments on the number of eminent men who were sons of carpenters or of farmers and the number of statesmen and ennobled aristocrats who were illegitimate children or the posterity of illegitimate children. (For further references see those cited in the pages of Trevelyan's *English Social History*, and for a more general review Fahlbeck, *Die Klassen und die Gesellschaft*.)

[1] His epitaph states that 'Wittington' "rose from indigence to wealth". Contemporary records, however, indicate that the original Sir Richard was in point of fact the son of a knight. Nevertheless, had social mobility not been a familiar possibility, these self-help legends, with their obvious morals, would never have become part of the native tradition.

factor at work must be some environmental handicap. But intelligence is by no means the sole characteristic of social importance: the presence or absence of certain motivational traits must also count. For the child to succeed, *both* high intelligence *and* high qualities of character are requisite; but *either* lack of intelligence *or* lack of those character-qualities may be sufficient to ensure a failure. This is evident from the case-studies we have carried out among pupils who have failed: those who showed high intelligence have commonly proved lacking either in certain moral traits (industry, perseverance, ambition, and the like) or in temperamental characteristics (readiness to undertake sedentary work with books and paper as contrasted with more active and practical work with people and things): all three are needed for an academic career.

The development of moral qualities even more than the development of intellectual qualities is largely dependent on environmental conditions. Nevertheless, we believe (though here the evidence is less conclusive) that differences in temperament and character are also in some degree the result of genetic tendencies. Indeed, in research on mental inheritance in animals it has proved easier to demonstrate the transmission of emotional and temperamental qualities than the transmission of cognitive efficiency. With rats, for example, it has been found possible by selective breeding to produce strains which appear innately ' bright ' or innately ' dull ' when tested by their performances in maze-learning; yet recent work has shown that the main difference between the two types lies not so much in their ' intelligence ', or their general capacity for learning, as in their emotional or motivational characteristics[1].

The Influence of Environment. Down to about twenty years ago the more ardent environmentalists attributed the lack of intelligence and of energy or drive, which seemed so marked in the ordinary pupil from the ' under-privileged groups ', to the poverty-stricken circumstances in which the unskilled or semi-skilled worker commonly lived[2]: his children were ill nourished and prone to

[1] C. S. Hall, Emotionality in the Rat, *J. Comp. Psychol.*, XXIV, 1938, 369–376 and refs. L. V. Searle, The Organization of Hereditary Maze-brightness and Maze-dullness, *Genet. Psychol. Mon.*, XXXIX, 1949, 279–325 and refs.

[2] No doubt, as Dr. Halsey and Dr. Floud maintain, in the past the wastage of talent through lack of opportunity for higher education was due to the financial handicaps of the ' under-privileged classes ' (by which they and other environmentalists mean the semi-skilled and unskilled workers). This they apparently suppose is the result of the unfair ' privilege ' enjoyed by the middle classes who continue to regard the grammar school as their ' undisputed preserve ' [12, p. 142]; and in spite of the alleged improvements " the probability that a working class boy will get to a grammar school is not (so they say) strikingly different from what it was " before the war. In Middlesbrough only " one working class boy in eight " gains admission to the grammar school as compared with " one in three " among pupils of " middle-class origin " (clerks and business people who would formerly have been in the fee paying class); in Hertfordshire the disproportion is still greater. However, when we look at the higher stages of education, one of the most serious causes of wastage today would seem to be the financial handicaps, not of the working, but of the middle class: e.g., under the heading of ' The Under-Privileged ' the editorial commentator of the *Times Educ. Sup.* (23 Jan. 1959) points out that " 2000 young people are kept out of universities through financial difficulties " and that " the victims are the children of the middle men "—i.e., of business people and professional workers: (detailed figures are quoted from the report of the Industrial Management Association).

a succession of epidemic diseases, and all this undermined both their ability and their character. Today, as Dr. Halsey and Dr. Floud acknowledge, such an explanation is no longer tenable. But they still maintain that differences in environment constitute essential factors; environment, so they now explain, must be conceived as operating " in more subtle forms ": a " general handicap is imposed on the child brought up in a working class family rather than in a middle class home background ". When we ask what exactly is the factual evidence for these ' subtler forms ' of environmental operation, we are merely told that " the precise nature of the hindrances urgently needs investigating ".

As a matter of fact, however, both in London and elsewhere, with the help of teachers, school medical officers, and care committee workers, a number of cooperative investigations have already been made into the material and the non-material (i.e. the ' subtler ') conditions in the environment which act as impediments to the child's educational progress (cf. *The Backward Child*, pp. 126f., 566f. and refs.). From the data so obtained it appears that such conditions unquestionably contribute both to the amount and to the degree of backwardness, dullness, and apparent feeble-mindedness that are so frequent among children from the lower social classes. In London unsatisfactory home conditions prove to be the major factor in about 13 per cent of the cases and a minor or contributory factor in just over half the cases. In about 8 per cent of the cases the main adverse factors were of a material nature—chiefly poverty and its concomitants: these have now decreased to less than 4 per cent. In 5 per cent they arose out of the unfavourable moral, emotional or cultural conditions in the home; and these non-material factors appeared as minor or contributory causes in over 22 per cent of the cases. In our studies of ' retarded ' cases among children with high I.Q.s—i.e. children whose innate ability was definitely above the average, but whose educational progress and whose performance in the scholarship or the 11 plus examination fell below their capabilities, the proportions were not quite so high. In all these groups the majority of those who suffered from environmental handicaps were children from class III (as defined above). Barely 10 per cent came from class I. And as a general rule, though by no means invariably, the child whose innate abilities are equal to or above the average appears able to surmount such obstacles without great difficulty, provided there is no serious defect of character (innate emotional instability, or the like), and provided he receives such assistance as a progressive school and an enlightened teacher will usually supply.

Now, although the material handicaps have largely been removed, the improvement in the psychological background has been much smaller; and, if these are the ' subtler conditions ' that Dr. Halsey and Dr. Floud have in mind, we should readily agree that they play an important part. Nevertheless, it is quite impossible to maintain, as they apparently would do, that such factors constitute the *sole* cause.

My final conclusion therefore is that none of the objections brought forward by Dr. Halsey suffice to refute either the deductive or the inductive arguments

I have described. Nor do I think he himself has been able to meet the criticisms urged by Miss Conway against his own hypothesis. Hence I think we are bound to accept the view that the differences in average ' intelligence ' exhibited by the different socio-economic classes in this country at the present time are mainly, though not perhaps entirely, the outcome of genetic difference.

V. Practical Corollaries

Educational Problems. The practical corollaries that follow from the general theory of mental inheritance adopted in this paper have been reviewed elsewhere ([6], pp. 81f and refs.). Here therefore I shall only touch on the more specific implications arising out of its application to the problem of class-structure. The first and most obvious is this. Any recent attempt to base our educational policy for the future on the assumption that there are no real differences, or at any rate no important differences, between the average intelligence of the different social classes, is not only bound to fail; it is likely to be fraught with disastrous consequences for the welfare of the nation as a whole, and at the same time to result in needless disappointments for the pupils concerned. The facts of genetic inequality, whether or not they conform to our personal wishes and ideals, are something that we cannot escape. On the other hand, the demand for able individuals is so pressing, and the supply so meagre, that we must do all we can to ascertain which particular children are endowed with high ability, and give them the fullest educational opportunities that their talents deserve, no matter from what social class they happen to spring. After all the differences between the means for the several classes are distinctly small when compared with the wide range of individual differences which (on the theory I have outlined) we must expect to find within each class.

Let us therefore inquire how far the actual process of selection, as reported in the investigations of Dr. Floud and her collaborators, complies with what we might reasonably require from an efficient procedure working in accordance with the hypothesis I have described. From earlier vocational studies we can make a fairly reliable assessment as to the average intelligence of the parents in each of the five classes specified by Dr. Floud and her colleagues in their book. From the multifactorial theory we can infer the probable intelligence of their children. We can express these assessments in terms of the conventional scale with a standard deviation of about 15 points. And from the correlation between the intelligence of children and their socio-economic class we can estimate the standard deviation within each of the five classes. As noted in my earlier inquiry, these theoretical values agree quite closely with the values obtained in our previous surveys; and rather surprisingly, as Miss Conway remarks, these earlier values are very similar to those reported by Dr. Floud and her colleagues[1].

[1] Dr. Floud and her colleagues apparently hold that children's performances in ' intelligence ' tests, of the type used in the 11 plus examinations in South-West Hertfordshire and Middlesbrough, are largely influenced by the education they have received. So far as the individual pupils and those particular tests are concerned, I should be inclined to agree. But it is not so easy to decide

At the time of their investigations, in the two areas studied the average percentage (out of the total age-group) who were successful in gaining entrance to a grammar school was just under 20 per cent. Such a figure would imply a borderline of 112·7 I.Q. We can now calculate the number of children from each of the five socio-economic classes whom we should expect, on the basis of the hypothesis I have outlined, to secure admission. These estimated percentages are shown in the top line of the table below. The actual percentages reported by the investigators themselves (averaged for the two areas) are set out in the second line of the table.

TABLE II. PERCENTAGE OF PUPILS FROM EACH SOCIAL CLASS ADMITTED TO A GRAMMAR SCHOOL

Class	1	2	3	4	5
Hypothetical	56·1	40·0	27·3	16·8	11·1
Actual	63·5	40·5	27·6	16·0	9·0

Description of classes (*Social Class and Educational Opportunity*, p. 28). 1. Professional workers, business owners, and managers. 2. Clerical workers. 3. Foremen, small shopkeepers, etc. 4. Skilled manual workers. 5. Unskilled manual workers.

It will be seen that the number of children from the working class who gained admission to a grammar school was indeed rather smaller than we should expect if innate ability was the sole criterion, while the number from the professional and clerical classes was appreciably higher. The discrepancies are greatest in the two extreme classes. To what then are the differences due ? Dr. Floud and Dr. Halsey would apparently attribute them to an injustice arising out of the use of so-called ' intelligence tests ': success at such tests, they believe, depend largely on ' acquired aptitudes ', and these ' aptitudes ' are not easily acquired in a working class home. This seems to me to be the least important of the causes. Rather more important are the differences in temperament and character: in explaining the causes of failure among London cases, ' laziness ', ' carelessness ', ' inattention ', ' lack of interest ', are far more frequently cited by the teacher in the case of bright children from the poorer classes than in the case of children from the professional or clerical classes. But by far the commonest reason is the child's poor performance in the formal tests of arithmetic and English: in the English test more especially children of ' unskilled labourers ' are undoubtedly penalized, whereas children from the ' professional ' classes derive a definite advantage from the intellectual background

in which direction such educational influences would operate. There is a good deal of evidence to suggest that, when the children have all been educated at primary schools under the same education authority, the effect on tests of a fairly mechanical type is to lessen the individual differences rather than create or augment them. Were that so, it might be argued that, if a test of ' pure ' intelligence could be applied, then the difference between the means for those coming from the lower and the higher social classes respectively would in fact be larger, and not (as Dr. Floud and her colleagues would have us believe) smaller or non-existent. However, my own guess would be that with tests of the Moray House type the opposing influences cancel out; and that seems to be confirmed by the agreement between the figures from these various inquiries.

of the cultured home. However, no matter what may be the cause there can to my mind be little question that the existing procedures still tend to miss some of the ablest pupils coming from the humbler classes, though the amount of failure is very much smaller than Dr. Halsey and Dr. Floud would have us believe.

Let me add that the foregoing deductions admittedly rest on a somewhat slender basis. But I have taken what seem to me the most reliable data available. The real need, however, is a fresh series of researches along the lines I have indicated, based not simply on the marks and measurements obtained in the 11 plus examination itself, but on a systematically planned set of inquiries carried out in a wide variety of typical areas.

Political Problems. The fact that the differences between the average intellectual level of the different socio-economic groups depends largely on genetic constitution raises a number of problems of social and national importance. The most obvious are those connected with the efficient organization of a democracy. Perhaps one of the clearest ways of bringing out the differences in what I will call ' effective intelligence ' is to consider the type of newspaper purchased by members of the different classes. An extremely thorough investigation of this problem is made at regular intervals by the National Readership Surveys of the Institute of Practitioners in Advertising. So far as occupational grouping is concerned their present classification of the population is as follows: *A*, Upper Middle Class (professional, higher administrative, etc.); *B*, Middle Class (intermediate professional or administrative); *C*1, Lower Middle Class (clerical, supervisory, etc.); *C*2, Skilled Manual Workers; *D*, Semi-skilled and Unskilled Manual Workers; *E*, The Poor (casual workers, widows, pensioners, etc., with under £6 per week). In the published tables the first two and last two classes are usually thrown together: thus we find that the daily morning paper most frequently read by classes *A*, *B*, and *C*1 is the *Daily Express*; that of *C*2, *D*, and *E*, the *Daily Mirror*; *The Times*, the *Observer*, and the *Sunday Times* are almost confined to class *A*; The *People* is the typical Sunday paper of class *C*2; the *News of the World* that of the classes *D* and *E*[1]. Almost equally informative are the findings of the various ' audience surveys ' relating to preferences for broadcast programmes and of several recent investigations into the type of advertisement addressed to members of each social class: (this may be readily inferred from those published in their favourite newspaper). The factors determining all these preferences are highly complex; and, although no adequate research has as yet been attempted on the psychological aspects of the problem, there can, I think, be little question that among all the various factors, differences in the average level of innate intelligence is one of the most important.

No doubt, there is still much truth in Walter Bagehot's epigram—that " public stupidity is the finest safeguard against the eccentricities of genius ".

[1] Cf. *National Readership Survey* [18]: the detailed tables are well worth studying.

Nevertheless, in a highly technological age the genius is equally needed to aid and supplement the comparative ineptitude of the general public; and, among nations, success in the struggle for survival is bound to depend more and more on the achievements of a small handful of individuals who are endowed by nature with outstanding gifts of ability and character. It is, and always has been, the ablest 2 or 3 per cent of the population to whom the great advancements of civilization are due—those who (in the technical jargon) must have had an I.Q. of roughly 130 or over[1]. Since their rate of development is 30 per cent greater than that of the average child, they will at 11 plus have a mental age of 15 or more. Hence it is obviously desirable that they should be sorted out for special education at the earliest possible age. Now, as we have seen, no method of discovery or selection can be perfect. We may, however, accept a correlation of 0·80 to 0·90 as expressing the predictive accuracy of the best procedures now available; and a reference to tables of the normal bivariate distribution will show that, in order to make reasonably sure of selecting the brightest 3 per cent for our grammar schools (or their equivalent), we must in actual fact place the borderline well below 130 I.Q. If we select the brightest 25 per cent as judged by the examination results (checked of course by the teacher's assessments), then we may be fairly certain that very few of the really able children—the cleverest of the clever—will be missed; probably not more than about 1 in 500 of the brightest 3 per cent would slip through the mesh. . This requires us to place the borderline at about 110 I.Q. The additional money spent on those moderately able children, who are, so to speak, of the second order of ability only, will not be wasted, since the ablest of all will need the cooperation of a band of intelligent co-workers to support their efforts.

But of course, as before, our selection will now involve a number of minor inaccuracies at this lower level: judged by their ultimate progress there will be an appreciable proportion among the rejected, i.e. among those allocated to the modern school or its equivalent, who will prove to have been quite as capable as many of those who were selected for the grammar school. But that is not of itself a reason for lowering the borderline still further, nor yet for discarding the whole scheme as inefficient or unjust: from a national standpoint the minor inaccuracies at these lower levels do not greatly matter; and, as has often been maintained, there are sound reasons why the lower socio-economic classes should not entirely be depleted of a few bright members here and there whose ability is well above the average[2].

[1] This figure would include the ablest section of our present university population (i.e. those who gain first-class honours degrees or are sufficiently equipped to carry out original research), together with their equivalents in industry. At the present time about 6 per cent of the total population reach the universities in this country. In America the figure is much higher (10 per cent); but even there the same slender proportion still holds good for those genuinely capable of original or first-class contributions.

[2] For simplicity I have expressed myself in terms of an educational scheme based on separate schools of the ' modern ' and ' grammar ' types. But this does not mean that I accept the popular notion that selection at 11 plus necessarily entails segregation in different buildings. The relative

If, however, our aim is to discover (so far as we are able) all those pupils who are genetically qualified to profit by a higher type of education, it will be a grave mistake to postpone our efforts until they reach the age of eleven: by then it is much too late to start estimating their innate capacities: the process of assessment should begin as soon as each child moves up from the infant school. And, when we have discovered them, it is essential to make the most of their educational potentialities. As things are, it is the average pupil in the class or school, and not the brightest, who tends to set the pace; and, in the hope of keeping the day-to-day progress more or less uniform, the teacher's extra efforts are devoted chiefly to pressing forward the weaker members: in consequence, as the results of educational tests clearly demonstrate, the brightest children are often, from the scholastic standpoint, the most retarded[1].

There is one further practical issue that arises out of the conclusions we have already reached. At the present time rapid changes are taking place not only in the traditional class-structure, but also (a point too often overlooked) in the traditional customs and ideals that have hitherto governed the mating habits prevalent in this country. Broadly speaking, we are changing from a system based mainly on inbreeding to one of relatively free outbreeding. Such a change, as has been amply demonstrated both by studies of mating systems in other organisms and by the past history of human civilizations, is commonly followed by striking modifications in the genetic constitution of the population concerned. Moreover, the situations thus produced are in effect undesigned experiments which, if followed up by impartial observation, may throw a much-needed light on the genetic principles at work below the surface. It is therefore urgently desirable that, before the relevant data pass from our grasp, systematic surveys and case-studies should be organized to investigate the effects of these changing conditions.

At present the policies of our government and ourselves are planned and carried out solely with an eye to the wishes and welfare of the present generation and with little or no regard to their possible effects on the generations still to come. In part this arises because the authorities who guide us are either sceptical or ignorant about the evidence for mental inheritance. For that

efficiency of ' comprehensive ' schemes as contrasted with bipartite or tripartite schemes is a distinct issue on which, in my view, the information at present available is insufficient for any definite conclusion to be reached (cf. [8]). Instead of heated debates or hasty decisions we need a free period for further experiments; and I strongly suspect that the final outcome will be that different schemes are suited to different areas. Here my object has merely been to bring together the main results emerging from a statistical study of existing data, and to indicate how they bear on current issues that are too often debated with no reference whatever to the actual facts.

[1] This is clearly brought out by calculating ' achievement ratios ' for each individual (not that I wish to recommend that particular formula). I fully agree that, since the brightest pupils in the modern school are of much the same intellectual capacity as the duller children in the grammar school, it is a sound and just policy to provide them, so far as possible, with a similar curriculum; nevertheless, when financial resources are limited, our desire to do all we can for the large mass of average children should not divert us from the still more urgent duty of doing our utmost for the ablest.

psychologists themselves are largely to blame[1]; and beyond all question the most effective remedy will be the prosecution and the publication of further research—research in which not only the geneticist but also the psychologist must in future play an active part.

References

[1] Bendix, R., and Lipset, S. M. (1954). *Class, Status, and Power.* London : Routledge and Kegan Paul.

[2] Burt, C. (1921). *Mental and Scholastic Tests.* King.

[3] Burt, C., *et al.* (1926). *A Study in Vocational Guidance.* London: H.M. Stationery Office.

[4] Burt, C. (1942). Ability and income. *Brit. J. Educ. Psychol.*, XIII, 83–98.

[5] Burt, C. (1946). Intelligence and fertility. *Papers on Eugenics*, No. 2. London: Hamish Hamilton.

[6] Burt, C. (1959). General ability and special aptitudes. *Educ. Research*, I, 81–96.

[7] Carr-Saunders, A. M., and Wilson, P. A. (1933). *The Professions.* Oxford: Clarendon Press.

[8] Conant, J. B. (1959). *The American High School Today.* New-York: McGraw-Hill.

[9] Conway, J. (1958). The inheritance of intelligence and its social implications. *Brit. J. Statist. Psychol.*, XI, 171–190.

[10] Darlington, C. D. (1958). The control of evolution in man. *Eugen. Rev.*, L. 113–123.

[11] Duff, J. F., and Thomson, G. H. (1924). The social and geographical distribution of intelligence in Northumberland. *Brit. J. Psychol.*, XIV, 162–173.

[12] Floud, J. E., Halsey, A. H., and Martin, F. M. (1956). *Social Class and Educational Opportunity.* London: Heinemann.

[13] Floud, J. E., and Halsey, A. H. (1958). Measured intelligence is largely an acquired characteristic. *Brit. J. Educ. Psychol.*, XXVIII, 290.

[14] Glass, D. V., ed. (1954). *Social Mobility in Britain.* London: Routledge and Kegan Paul.

[15] Gretton, R. H. (1917). *The English Middle Class.* London: Bell.

[16] Halsey, A. H. (1958). Genetics, social structure, and intelligence. *Brit. J. Sociol.*, IX, 15–28.

[17] Hogben, L., ed. (1938). *Political Arithmetic.* London: Allen and Unwin.

[18] Institute of Practitioners in Advertising (1954–8). *National Readership Survey.* London: Inst. Pract. Adv.

[19] Maddox, H. (1958). Nature-nurture balance-sheets. *Brit. J. Educ. Psychol.*, XXVII, 166–175.

[20] Mannheim, K. (1940). *Man and Society.* Kegan Paul, Trench, Trubner.

[21] Warburton, F. W. (1958). A reply to ' Measured intelligence is largely an acquired characteristic '. *Brit. J. Educ. Psychol.*, XXVIII, 291–292.

[1] It is unfortunate that so many writers in this country still cite the findings of American pre-war investigations as though they provided the final word on the matter. Thus, as Dr. Maddox has recently pointed out, " the last major American ' nature-nurture ' survey—that published in the *Thirty-ninth Yearbook* (1940)—only succeeded in revealing the apparently irreconcilable differences between traditionalists (i.e. hereditarians) such as Terman and Goodenough and Stoddard and Wellman ". And Dr. Maddox himself appears to endorse the inference drawn by one eminent contributor to that volume, Professor Harold E. Jones, namely, that " potential contributors should now be advised that nature-nurture has now been shown to be an unproductive field of research " [19, p. 154]. Few American psychologists today, I fancy, would still accept at their face value the conclusions reached by Stoddard and Wellman twenty years ago; and certainly no geneticist would support Dr. Jones' advice. For the geneticist's view cf. Darlington [10].

Vol. XII Part I The British Journal of Statistical Psychology May 1959

THE EFFECT UPON THE COMMUNALITIES OF CHANGING THE ESTIMATE OF THE NUMBER OF FACTORS

By Charles Wrigley[1]
Michigan State University

This paper is concerned with the arithmetical consequences of adopting the theory of minimal rank as a means of determining communalities in factor analysis. The complexities and inconsistencies inherent in the algebraic formulation are illustrated by an analysis of Burt's 11×11 table of correlations between emotional traits. It is found that, owing to the appearance of Heywood cases, the only solution that meets Thurstone's requirements entails more factors and higher communalities than the minimal rank would imply, and that this solution has eventually to be rejected on psychological grounds. The final conclusion is that, even were a satisfactory method found for determining the communalities implied by minimal rank, the theory would still be unacceptable because it neglects the statistical considerations imposed by the fact that every empirical matrix is obtained by sampling.

I. Introduction

According to Thurstone the problem of communalities consists in finding those diagonal values which will minimize the rank of the correlation matrix, and thereby account for the observed coefficients in terms of the minimal number of common factors. "The relations between the correlations and the unknown communalities", he writes, "can be stated in the theorem that *the smallest number of independent factors that will account for the intercorrelations of* n *tests is the minimum rank of the correlation matrix with the diagonal entries treated as unknown positive values between* 0 *and* +1" ([15], p. 283, his italics). Thus the practical problem of deciding on (i) the number of factors and (ii) the values for the communalities are both made to depend on the algebra of minimal rank.

In the case of empirical data, however, Thurstone introduced an important qualification. "The scientific problem of analysing a set of experimentally given correlation coefficients is to find another correlation matrix of minimum rank whose discrepancies from the experimentally given coefficients are small enough to be ignored" ([15], p. 283): that is to say, he proposes factoring, not until the residuals are zero, but only until they are very small, since sampling errors make it unprofitable to fit correlations to empirical data with absolute exactitude.

In *Multiple Factor Analysis* Thurstone presents no test to determine when the residuals are small enough for the factoring to stop; nor does he refer

[1] I am indebted to Professor L. G. Henyey (of the University of California Computer Center) for allowing use of the IBM 701 electronic computer in this research and to Mr. J. O. Neuhaus for preparing the programme for the iterative estimation of the communalities. The work of the Computer Center has been aided by a grant from the National Science Foundation.

to Lawley's maximum likelihood method of factoring [11], which provides a test for the significance of the factors; instead he leaves the decision as to the number of factors to the investigator's judgement: "it is a question of statistical judgement when the residuals are small enough to be ignored" ([15], p. 295). This means that different investigators, starting from the same data, will extract a different number of factors, and obtain different communalities. But once this concession about ignoring small residuals has been introduced the minimal rank requirement no longer plays the main functional role in determining the communalities. Since the number of factors now depends on the good judgement of the investigator, there is no objective basis for preferring one solution to another. Empirical investigation becomes restricted to the narrower problem of finding how much communalities change with any increase in the estimate of the factors.

The problem can be recast in objective terms by returning to a strict insistence on minimal rank, i.e. by disregarding statistical considerations, and requiring factoring to continue until all residuals are zero. (There remains the possibility that minimal rank can be achieved by more than one set of communalities within the range of 0 and +1, so that some further restriction may have to be introduced to ensure a unique solution.) When Albert [1] examined the problem of communalities, he interpreted the minimal rank requirement in this strict way. But the practical utility of such an approach is limited, because no procedure is supplied for translating the general principle into arithmetical practice. No serviceable method is known for a direct determination of minimal rank. Thurstone had to advocate an iterative procedure: the factoring begins with trial values for the communalities, and proceeds until residuals are very small. Results of the provisional solution are then used to provide better estimates of the communalities, and so of the number of factors. It will be noticed that the desideratum of minimal rank once again plays a non-functional role. While the logic makes the communalities dependent on the rank, it is only after the communalities are known that the minimal rank can be determined. The investigator is thus faced with a vicious circle.

The aim of this paper is to examine the arithmetical consequences of this interlocking of the communalities with the number of common factors. Section II will review some of the mathematical findings upon rank reduction. Section III describes the procedure followed in the empirical investigation reported in this paper. Section IV gives the results; and Section V discusses their main implications.

II. The Mathematics of Rank Reduction

The algebraic findings on rank reduction have recently been reviewed by Kaiser ([10], pp. 4–14); so that no comprehensive review need be attempted here. This summary will be confined to certain findings and unsettled problems relevant to the research that follows.

The Determination of Minimal Rank. While no serviceable procedure has so far been developed for the direct determination of minimal rank, lower and

upper bounds have been supplied. Guttman [5] proved the lower bound to be given by the number of positive latent roots when the squared multiple correlations of each of the p variables with the remaining $p-1$ variables are inserted in the leading diagonal of the correlation matrix. Thurstone ([15], p. 308) proved that every correlation matrix may be expressed in terms of not more than $p-1$ common factors, and that there will be at least one set of diagonal values within the range of 0 and $+1$ which will do so. He reached this result by considering the fact that, in the square root method of factoring, there remains only one non-zero residual after the extraction of $p-1$ factors, and that this, being a diagonal entry, may be attributed to test specificity.

It is not difficult to show that there is an infinite number of ways in which the rank of a matrix can be reduced to $p-1$. Let any constant q be subtracted from all elements of the leading diagonal of a square symmetric matrix A. If the latent roots of A be denoted by $\lambda_1, \ldots, \lambda_p$, then the latent roots of $[A-qI]$ will be $\lambda_1-q, \ldots, \lambda_p-q$ ([3], III, p. 309). If q be set equal to λ_p, then the smallest root of $[A-qI]$ will be zero[1]. Now let $k_1, \ldots, k_p$ be the diagonal values for which the correlation matrix R is positive definite: (this generally holds with unities in the diagonal). Rank reduction will then be achieved with the diagonal values $k_1-\lambda_p, \ldots, k_p-\lambda_p$. Add any amount δ to k_1. All latent roots will then be increased. Let $\lambda_p+\gamma$ be the smallest latent root in the new matrix. Rank reduction to $p-1$ can be achieved with the new set of diagonal values

$$k_1+\delta-\lambda_p-\gamma, \quad k_2-\lambda_p-\gamma, \ldots, k_p-\lambda_p-\gamma.$$

Ledermann [12] established the possibility of writing a set of equations to find communalities which will reduce rank to $[2p+1-\sqrt{(8p+1)}]/2$; but there is no guarantee that (*a*) the equations will be solvable (see Guttman's discussion [7]), or (*b*) if they are, that the communalities will lie within 0 and $+1$.

The Problem of Non-uniqueness of Communalities. Thurstone made no allowance for the possibility that more than one set of communalities might meet the minimal rank requirement. He writes that he "assumed, in the development of multiple-factor analysis, that there exists a unique set of communalities consistent with the experimentally given off-diagonal correlations" ([15], p. 318). This has been found to hold for matrices of small rank, but not otherwise. Three situations can be distinguished:

1. Matrices of minimal rank of less than $p/2$. Albert [1] proved that the communalities will then be unique.

2. Matrices of minimal rank greater than $[2p+1-\sqrt{(8p+1)}]/2$. Ledermann [12] proved that there will then be an infinite number of solutions yielding minimal rank: Thomson ([14], pp. 81–82) provides an example.

3. Matrices of minimal rank equal to or greater than $p/2$, but not exceeding $[2p+1-\sqrt{(8p+1)}]/2$. Wilson and Worcester [17, 18] established the existence

[1] If there happen to be multiple roots equal to λ_p, the rank will be reduced by more than one.

of multiple solutions, the number depending on the minimal number of communalities that must be included in a minor of the order that is one greater than the minimal rank: Thurstone ([15], p. 310) supplies an example. These algebraic results relate to any diagonal values; and the communalities will not necessarily lie within 0 and +1.

What happens when communalities are required either for fewer or for more factors than the minimal rank number the position is not entirely clear. When an attempt is made to reduce rank below the minimum by the standard iterative procedure, it seems to be assumed (i) that a least-squares solution will be obtained in which the sum of squares of residuals is minimal for the given number of factors (namely, a sum equal to that of squares of the rejected latent roots) and (ii) that only one set of communalities will achieve this. If the rank to be fitted is greater than the minimum, then it is apparently supposed that a new set of communalities can be found to fit to a rank which is one greater than the minimum, then another set to fit a rank which is two greater than the minimum, and so on. Thurstone ([15], pp. 289–290) describes this as appropriate for the Heywood case, and indicates that there is an infinite number of sets which might be found in reducing a rank greater than the minimum.

III. Arithmetical Procedure

In order to illustrate and investigate the arithmetical consequences of the method, an eleven-variable matrix, reported and analysed by Burt in an earlier issue of this journal ([2], pp. 199 f), was selected for study. The matrix consists of an 11×11 table of observed correlations between assessments for the chief 'primary emotions', obtained for the purposes of a factorial research on temperamental characteristics[1]. The observed correlations are set out in Table I: the values shown (within parentheses) in the leading diagonal are the Squared Multiple Correlations (SMC's) of each variable with the remaining $p-1$ variables. The same correlation matrix has already been used in an investigation into the efficiency of various methods for estimating communalities [19]. In that investigation it was shown that the Squared Multiple Correlations would provide effective estimates with which to start the iterative procedure, but that zeros in the diagonal furnish a good starting-point when the iteration is prolonged. In this earlier study, however, communalities were calculated for seven common factors only. In the present research the estimated number of common factors ranges from one to ten. In one series zeros are used as initial estimates, in another SMCs, and for certain estimates unities are used instead.

The iteration started with a calculation of principal axes factor loadings, based on whichever initial estimates were chosen for the communalities. The sum of squares of the loadings for the estimated number of common factors was

[1] The research was part of an investigation reported in a paper on 'General and Specific Factors underlying the Primary Emotions' (*Brit. Ass. Ann. Rep.*, 1915, pp. 695 f.). Further details are given in the Appendix to this paper.

TABLE I. MATRIX OF OBSERVED CORRELATIONS

Trait	1	2	3	4	5	6	7	8	9	10	11
1. Sociability	(0·927)	0·821	0·703	0·754	0·723	0·732	0·538	0·569	0·540	0·294	0·117
2. Sorrow	0·821	(0·859)	0·804	0·583	0·561	0·443	0·522	0·400	0·249	0·487	0·282
3. Tenderness	0·703	0·804	(0·783)	0·645	0·372	0·316	0·308	0·142	0·241	0·401	0·342
4. Joy	0·754	0·583	0·645	(0·660)	0·506	0·502	0·343	0·306	0·464	0·181	−0·016
5. Wonder	0·723	0·561	0·372	0·506	(0·606)	0·534	0·377	0·547	0·453	0·108	−0·133
6. Elation	0·732	0·443	0·316	0·502	0·534	(0·652)	0·469	0·506	0·469	0·113	0·018
7. Disgust	0·538	0·522	0·308	0·343	0·377	0·469	(0·378)	0·384	0·251	0·227	0·084
8. Anger	0·569	0·400	0·142	0·306	0·547	0·506	0·384	(0·495)	0·419	0·038	−0·110
9. Sex	0·540	0·249	0·241	0·464	0·453	0·469	0·251	0·419	(0·433)	−0·048	−0·155
10. Fear	0·294	0·487	0·401	0·181	0·108	0·113	0·227	0·038	−0·048	(0·304)	0·304
11. Subjection	0·117	0·282	0·342	−0·016	−0·133	0·018	0·084	−0·110	−0·155	0·304	(0·298)
Latent Roots	4·795	1·330	0·443	0·213	0·047	0·001	−0·038	−0·066	−0·081	−0·113	−0·136

then taken as the revised communality estimate, and new loadings obtained by a new principal axes analysis. In each series iteration was continued until the successive communality estimates varied only slightly from the preceding estimates. With an IBM 701 electronic computer about 45 minutes were needed for each set of a hundred iterations.

IV. Results

Results will be reported in the following order: (1) the effect on the communalities of increasing the estimated number of factors; (2) the determination of the minimal rank; (3) the rate of convergence of the iterative procedure with different numbers of factors; (4) the non-uniqueness of the communalities for the larger ranks; (5) a comparison of the starting-points for iteration with different numbers of factors; (6) the more rapid convergence of the more extreme estimates of communalities.

(1) *Effect on Communalities of Increasing the Estimated Number of Factors.* Estimates of communalities at the conclusion of the iterations for one to ten factors are shown in Table II[1]. These were obtained with zeros as the starting point. Latent roots (i.e. sums of squares of loadings) for the successive factors are shown in Table III[1].

Table II. Communalities for Different Estimates of the Number of Factors (Starting Point, Zeros)

	Number of Factors									
	1	2	3	4	5	6	7	8	9	10
Number of Iterations	20	25	48	89	95	100	151	94	138	115
1	1·041	1·010	0·993	0·992	0·991	1·008	1·004	1·002	1·003	1·004
2	0·697	0·893	0·902	1·029	1·028	1·020	0·999	1·011	1·009	1·007
3	0·463	0·742	0·916	0·889	0·874	0·878	0·946	0·922	0·909	0·906
4	0·539	0·528	0·697	0·708	0·739	0·735	0·761	0·757	0·754	0·749
5	0·498	0·591	0·584	0·636	0·627	0·653	0·678	0·691	0·681	0·676
6	0·470	0·545	0·553	0·697	0·682	0·684	0·698	0·704	0·708	0·707
7	0·314	0·308	0·384	0·378	0·431	0·470	0·475	0·466	0·461	0·461
8	0·298	0·449	0·547	0·542	0·580	0·592	0·595	0·596	0·599	0·600
9	0·256	0·422	0·458	0·453	0·462	0·493	0·480	0·481	0·485	0·484
10	0·096	0·314	0·347	0·325	0·328	0·325	0·338	0·349	0·346	0·344
11	0·011	0·290	0·304	0·414	0·622	0·599	0·521	0·519	0·537	0·547
Sum	4·683	6·092	6·685	7·063	7·364	7·457	7·495	7·498	7·492	7·485

[1] The results in both these tables are rounded to three decimal places : they were printed by the computer to eight places.

TABLE III. LATENT ROOTS FOR DIFFERENT ESTIMATES OF THE NUMBER OF FACTORS

(STARTING POINT, ZEROS)

	Number of Factors									
	1	2	3	4	5	6	7	8	9	10
Number of Iterations	19	24	47	88	94	99	150	93	137	114
1	4·686	4·778	4·828	4·865	4·871	4·881	4·891	4·891	4·889	4·887
2	1·106	1·313	1·370	1·408	1·461	1·460	1·454	1·452	1·454	1·456
3	0·283	0·380	0·487	0·499	0·535	0·541	0·554	0·548	0·546	0·545
4	0·060	0·179	0·190	0·291	0·346	0·341	0·320	0·325	0·329	0·331
5	0·007	0·011	0·082	0·099	0·152	0·167	0·165	0·161	0·162	0·163
6	−0·128	−0·008	0·028	0·037	0·047	0·068	0·076	0·077	0·075	0·074
7	−0·164	−0·042	−0·001	0·006	0·008	0·011	0·037	0·032	0·028	0·026
8	−0·230	−0·062	−0·044	−0·010	−0·001	0·001	0·003	0·009	0·007	0·005
9	−0·245	−0·127	−0·050	−0·028	−0·014	−0·000	0·000	0·000	0·001	−0·000
10	−0·297	−0·137	−0·077	−0·044	−0·018	−0·003	−0·001	0·000	0·000	−0·000
11	−0·392	−0·194	−0·128	−0·059	−0·022	−0·008	−0·002	−0·000	−0·000	−0·000

Thurstone apparently assumed that the introduction of further factors would lead to an increase in the communalities: "since each factor contributes a small increment to the communality, it is evident that the estimated communality must be judged in relation to the number of factors that are extracted" ([15], p. 295). Usually the communalities increase, but they do not always do so. There are three that decrease in passing from one to two factors; there are two decreases in passing to three; six in passing to four; and so on. The irregularity is demonstrated still more clearly by considering a wider span of factors. For convenience let us restrict attention to the three-factor, six-factor, and nine-factor solutions. The communalities (variables 3 and 10) decrease in increasing from three to six: (even with the nine-factor solution, the communalities for these two variables remain less than they were for three-factors). In increasing from six to nine factors, five communalities decrease: variable 11 shows particularly irregular estimates: the communality is 0·01 for one factor, 0·29 for two, and 0·62 for five, but declined to 0·52 for eight. To decide with any precision what is the amount of the common variance in variable 11 thus seems almost impossible: on the whole the latent roots are more regular, but even here there are occasional decreases.

(2) *What is the Minimal Rank of the Matrix*? Thurstone's requirement of minimal rank with diagonal entries between 0 and +1 is met by none of the ten solutions. The lower bound for the rank of the matrix must be six, since there are six positive latent roots when SMC's are inserted in the diagonal. In all results with six factors or more, at least one, sometimes two of the communalities prove to be larger than +1. In general the evidence does not allow us to decide whether this weakness arises from the minimal rank theory or from the iterative procedure. In the special case of rank ten, however, the failure clearly lies in the iterative procedure. With unities in the diagonal, the smallest latent root is 0·056, so that the rank can be reduced to ten by setting all communalities equal to 0·944; and this solution avoids a Heywood case. Iteration from zeros, however, converges neither to this nor to any other acceptable solution, but to an unacceptable one with two communalities greater than +1.

The trace, i.e. the sum of the communalities, is for the iterative solution 7·485, whereas for the communalities of $1-\lambda_p$ it is 10·384. Thus the solution with the smaller trace results in a Heywood case, but with larger trace this result is avoided. This particular matrix was selected for study because of the excellence of Burt's own communality estimates, as measured by the sum of squares of rejected roots: (cf. [19], pp. 21–22)[1]. At the time of the selection, the sampling details were not available. In the paper cited, however, Burt ([2], p. 198) states that the correlations were "the averages obtained by pooling the four largest groups" of subjects, "numbering 287 children in all". The persistence of the Heywood case is possibly attributable to this, so that further correlation matrices should be investigated. The present evidence allows only the qualified conclusions: (*a*) that the Heywood case will possibly be encountered in day-to-day

[1] Cf. Appendix, p. 52, Table X.

practice, and (*b*) that, when such a case exists, it may not always be resolvable by the relatively simple expedient of increasing rank by one.

(3) *Rate of Convergence with Different Estimates for the Number of Factors.* Stability of iteration at any stage can be conveniently measured by summing squares of differences in communalities from one iteration to the next. These values are reported in Table IV for the tenth, fortieth, and eightieth iterations. Convergence is fastest with one factor only, and becomes progressively slower with an increasing number, until the attempt is made to reduce to rank six, which seems to offer the most complex situation. For one factor, stability (to eight decimal places) is achieved after about twenty iterations; for two factors, after thirty iterations; for three factors, after fifty iterations; and for four factors, after about eighty iterations. Thereafter communalities continue to change in later decimal places even after eighty iterations. But the figures are not entirely regular. The estimates for rank nine, for example, although changing very markedly at first, eventually stabilize before those for either rank eight or rank ten.

Table IV. Communality Changes for Different Estimates of Rank (Starting Point, Zeros)

Estimate of Rank	Stage in Iteration 9–10	39–40	79–80
1	109	0	0
2	644	0	0
3	147274	5	0
4	630659	50	0
5	1267509	39518	10052
6	1166325	28680	9137
7	1950256	3999	582
8	7968511	1496	35
9	41599848	830	0
10	129050227	536317	119

Note. Communality changes are here measured by the sum of squares. Entries in the table have been multiplied by 10^{10}.

(4) *The Non-Uniqueness of Communalities for Larger Rank Estimates.* We saw that multiple solutions for communalities exist when the rank is $p/2$ or greater. An important practical problem is whether the iterative procedure from different initial estimates converges to different sets, or always to the same set, no matter what may be the postulated number of factors. This was examined in the earlier rank seven study ([19], pp. 14–16). Convergence was too slow for communalities to become completely stable. Hence it was hard to decide whether discrepancies represented non-uniqueness or merely a too early termination of the iteration. There had to be the qualified conclusion: "at least from a practical viewpoint

it appears possible to arrive at very different estimates each giving a very good fit to the correlations" ([19], p. 16).

The relatively rapid convergence to rank nine makes it more suitable for studying this problem. Reduction of rank, to at least six decimal places, is obtained after about 150 iterations. Results are unequivocal. Quite different communalities are obtained by iterating from unities, and from zeros, or SMC's. As will be seen from Table V, the communalities derived by iterating from unities do not conform to psychological expectations. Variables 4 and 9 now have communalities greater than +1, whereas previously this happened only with variables 1 and 2. Variable 11, with an SMC of only 0·298, has a communality of almost +1. On the other hand, variable 3, with the third largest SMC, has the lowest (the only one less than 0·900).

TABLE V. ALTERNATIVE SOLUTIONS FOR RANK NINE

	Starting Point					
Iterations	−3s	Zeros	Averages*	SMC's	Unities	+3s
1	1·003	1·003	1·001	1·000	0·940	0·836
2	1·009	1·009	1·008	1·006	0·929	0·919
3	0·910	0·909	0·918	0·927	0·824	0·742
4	0·754	0·754	0·760	0·767	1·001	2·903
5	0·680	0·681	0·693	0·706	0·985	2·841
6	0·707	0·708	0·713	0·717	0·994	2·832
7	0·462	0·461	0·457	0·452	0·994	2·926
8	0·599	0·599	0·594	0·591	0·986	2·883
9	0·485	0·485	0·487	0·490	1·001	2·970
10	0·345	0·346	0·351	0·357	0·998	2·967
11	0·537	0·537	0·518	0·500	1·000	2·997
Sum	7·491	7·492	7·500	7·513	10·652	25·816

* The starting points for this series were the averages of the communalities derived from the zero and from the SMC series.

Furthermore, the communalities obtained by starting from zeros differ from those obtained from SMC's, though to a much smaller extent (0·037 is the largest difference). A fourth set of communalities was obtained by averaging the communalities derived from the zeros series and the SMC series. To demonstrate even more extreme ways of reducing rank, i.e. starting with diagonal values quite unlike communalities as normally understood, iteration was started from arbitrarily high and low diagonal values. The first series started with +3 in each diagonal element, the second series with −3. The high initial values achieve rank reduction with eight communalities exceeding +2, but two others (those for variables 2 and 3) less than the SMC[1]. The low initial values provide communalities like those obtained in the zero series, though not exactly the same: the sum is 0·0006 lower. For one variable the difference is more than 0·001,

[1] Guttman [4] has proved that no communality can be less than the SMC ; but his proof requires that no communality exceed +1.

and for two others exceeds 0·0005. Values were calculated to eight decimal places, so that these are certainly real differences and not due to rounding. (Calculations start afresh from the correlations in each iteration, so that rounding errors do not accumulate.)

With non-uniqueness established for rank nine, further consideration was given to the rank seven situation, including the use of extreme values such as +3 and −3, and further iteration with the zero, SMC, and unity series. Results are presented in Table VI. Because of slow convergence, no iteration was completed. The evidence suggests, however, that non-uniqueness exists for rank seven as well as for rank nine. Finally, communalities were obtained for rank two with zero, SMC, and unity values in the diagonal. The same communalities were then obtained in every case. This strongly confirms the belief that the communalities are unique when a least squares solution is obtained for fewer factors than minimal rank.

TABLE VI. ALTERNATIVE SOLUTIONS FOR RANK SEVEN

	Starting Point				
Iterations	−3s 93	Zeros 151	SMC's 158	Unities 204	+3s 108
1	1·003	1·004	1·003	1·012	0·950
2	0·999	0·999	0·998	0·945	0·811
3	0·939	0·946	0·946	0·900	0·755
4	0·761	0·761	0·758	0·700	0·592
5	0·681	0·678	0·681	0·626	0·506
6	0·705	0·698	0·716	0·648	2·517
7	0·466	0·475	0·458	0·978	2·901
8	0·597	0·595	0·595	0·563	1·370
9	0·479	0·480	0·481	0·926	2·771
10	0·339	0·338	0·339	0·975	2·883
11	0·520	0·521	0·524	0·777	2·556
Sum	7·489	7·495	7·499	9·050	18·612
Square-sum of rejected roots	15	13	14	227	42921

NOTE. The values for the square-sum have been multiplied by 10^6.

(5) *Zeros or SMC's as Starting Points for Iteration* ? Because of the non-uniqueness of the communalities, some rule is needed for selecting one set rather than another. The psychometric approach to factorial theory defines the communality as a measure of the theoretical common variance in a domain of tests with factorial composition similar to those in the sample, as Guttman [5] and Tryon [16] have shown, whereas the SMC measures the observed common variance in the sample. It seems reasonable therefore to accept the set of communalities which agrees most closely with the SMC's, i.e. more theoretical

measures of common variance which agree best with the observed common variance. Since the SMC is the lower bound for the communality (when the communality is associated with rank reduction), the set with minimal trace will be that which agrees best with the SMC's.

The communalities problem is unfortunately no more easily solved for minimal trace than for minimal rank. It would be useful if a zero series invariably led to minimal trace; but the current study provides at least two counter-examples. We have already seen that the rank solution with the arbitrarily low values of -3 has a slightly smaller trace than the corresponding zero solution. Moreover, the trace for a zero series is not necessarily less than for a corresponding SMC series (when calculated for the same number of iterations), although this is usually the case. Table VII presents communalities for the SMC series with ranks one to ten. For rank six, the trace for the zero series is approximately 0·00012 greater for the zero series than for the SMC series (100 iterations).

TABLE VII. COMMUNALITIES FOR DIFFERENT ESTIMATIONS OF THE NUMBER OF FACTORS (STARTING POINT, SMC's)

	Number of Factors									
	1	2	3	4	5	6	7	8	9	10
Iterations	20	32	46	108	80	100	158	96	150	62
1	1·041	1·010	0·993	0·992	0·991	1·007	1·003	1·001	1·000	1·006
2	0·697	0·893	0·902	1·029	1·028	1·020	0·998	1·006	1·006	1·001
3	0·463	0·742	0·916	0·889	0·875	0·878	0·946	0·933	0·927	0·929
4	0·539	0·528	0·697	0·708	0·741	0·736	0·758	0·761	0·767	0·776
5	0·498	0·591	0·583	0·636	0·628	0·654	0·681	0·706	0·706	0·707
6	0·470	0·545	0·553	0·697	0·682	0·686	0·716	0·715	0·717	0·716
7	0·314	0·308	0·384	0·378	0·432	0·469	0·458	0·454	0·452	0·454
8	0·298	0·449	0·547	0·542	0·579	0·592	0·595	0·591	0·591	0·594
9	0·256	0·422	0·458	0·453	0·462	0·493	0·481	0·483	0·490	0·490
10	0·096	0·314	0·347	0·325	0·328	0·325	0·339	0·356	0·357	0·361
11	0·011	0·290	0·304	0·414	0·611	0·596	0·524	0·500	0·500	0·492
Sum	4·683	6·092	6·684	7·063	7·357	7·456	7·499	7·506	7·513	7·526

At present therefore the only practical value of such a minimal trace condition is (*a*) to indicate that iteration should begin with zeros as more likely to provide a small trace than any other initial values so far investigated, and (*b*) to enable us to select the solution with the smallest trace, when alternatives are available for the same rank.

Differences between the communalities obtained from the zero and the SMC series increase with the increase in the number of factors. Possibly, as rank is increased, there is a wider range of possible values for the communalities: (this result may be partly due to less complete convergence in solutions with more factors; but this cannot be the sole reason since convergence tends to speed up

again after rank six). Fortunately, even when the number of factors is large, the differences are so small that the psychological description of test specificity will hardly be changed whether iteration starts with zeros or with SMC's.

A much greater difference is introduced when unities form the starting point. Table VIII shows the results for ranks two, five, seven, and nine. They agree with those for the zero series in the case of rank two and do not differ greatly in the case of rank five (a trace of 7·578 instead of 7·364). They are much larger, and follow an entirely different pattern, with rank seven (9·050 instead of 7·495) and rank nine (10·652 instead of 7·492). Moreover, the communalities obtained in this way do not agree nearly so well with the expectations based on the correlations and the SMC's. The evidence therefore seems conclusive that more reasonable communalities (i.e. values agreeing more closely with the SMCs) will be obtained by iterating from zero or from SMC's.

TABLE VIII. COMMUNALITIES FOR DIFFERENT ESTIMATIONS OF THE NUMBER OF FACTORS (STARTING POINT, UNITIES)

	Number of Factors			
	2	5	7	9
Iterations	30	132	204	100
1	1·010	0·998	1·012	0·940
2	0·893	1·036	0·945	0·929
3	0·742	0·849	0·900	0·824
4	0·528	0·742	0·700	1·001
5	0·591	0·625	0·626	0·985
6	0·545	0·614	0·648	0·994
7	0·308	0·686	0·978	0·994
8	0·449	0·570	0·563	0·986
9	0·422	0·458	0·926	1·001
10	0·314	0·306	0·975	0·998
11	0·290	0·694	0·777	1·000
Sum	6·092	7·578	9·050	10·652

(6) *More Rapid Convergence from More Extreme Initial Estimates.* In the earlier study, convergence was found to be more rapid for low initial estimates than for high ([19], pp. 16–17). The initially poorer estimation provided by zeros, for example, is generally redressed by more rapid convergence than for the SMC series. This finding is maintained in the present investigation. Up to rank seven, the zero series generally overtakes the SMC series within about 20 or 25 iterations. For rank eight, about 35 iterations are needed; for rank nine, nearly 80; and with rank ten, the zero series never overtakes the SMC one. On balance therefore the advantage in iteration seems to rest with zeros as starting point.

The present study, however, suggests that the advantage of having low estimates diminishes as higher estimates of rank are reached. Our conclusion therefore should have been that convergence is more rapid for more extreme estimates than for the less extreme. The results for rank nine (Table IX) supply an example. The zero series overtakes the SMC series; and the +3 series overtakes the unities series. But the high extreme values converge more rapidly than the low extreme values, so that the +3 series converges more rapidly than any other. However, we have already seen that communalities obtained with these high values are quite unacceptable on other grounds.

TABLE IX. SPEED OF CONVERGENCE FOR RANK NINE

Number of Iterations needed to attain a Specified Degree of Accuracy

	Initial Estimates				
Largest Rejected Latent Root	−3s	Zeros	SMC's	Unities	+3s
10^{-3}	45	38	31	30	32
10^{-4}	65	58	55	49	41
10^{-5}	85	79	80	69	51
10^{-6}	108	102	106	89	60
10^{-7}	133	126	132	—	69

V. GENERAL CONCLUSIONS

The highly complex arithmetic of minimal rank is evident from these results. For more than twenty years the theory of communalities has been expressed in terms of minimal rank. The results of the foregoing study suggest that the stress upon the algebra of minimal rank may have served to complicate rather than clarify the problem. Four major objections may be raised against this purely algebraic approach: (*a*) the difficulty of finding minimal rank (with communalities between 0 and +1); (*b*) the existence of alternative solutions; (*c*) the unacceptability on psychological grounds of a minimal rank solution even if it can be attained; (*d*) the neglect of statistical considerations. Each of these will now be examined in turn.

(*a*) *The Difficulty of Finding Minimal Rank.* The correlation matrix in this research has been subjected to intensive analysis, far more than would ever be practicable in an ordinary day-to-day analysis, yet it has proved quite impossible to determine the minimal rank. Certainly the lower bound is definitely six. Is the minimal rank also six ? After 100 iterations the five smallest roots for rank six range from 0·011 to −0·008; the rank therefore seems likely to be more than six, although, since convergence is not complete, one cannot be quite certain. The rejected roots might prove in the end to be zero. For rank seven, the four smallest roots range (after 150 iterations) from 0·003 to −0·002. And there is the same dilemma as before. Ledermann [12], it is true, has shown that equations could be written to provide for a reduction to rank seven; but this is not conclusive since the equations are not necessarily soluble [7].

We can be fairly certain that reduction to rank eight is possible, since the last three roots are zero, when rounded to three places. But in both the zero and the SMC solutions, two communalities are greater than +1, so that the results are unacceptable psychologically. For rank nine, the further roots become exactly zero (within the limits of machine calculation), but there still remain two communalities greater than +1. The Heywood case remained for rank ten when the standard iterative procedure was used: this might possibly have been avoided by not allowing any diagonal value to exceed +1 during iteration; but it is clear that the algebra of rank will not necessarily fit our psychological preconceptions. And the only solution which meets Thurstone's requirements is that for rank ten with all communalities equal to 0·944; and these communalities are psychologically unreasonable because of the wide variations in the SMCs.

(*b*) *The Problem of Non-uniqueness.* The precise relations necessary to insure a rank of less than $p/2$ would seem to be extremely unlikely to be fulfilled with any empirical set of correlations. Once the rank is $p/2$ or greater, there is the ugly and insufficiently considered problem of alternative sets of communalities. What are the psychological grounds for selecting one solution as informative about the proportion of common and specific variance, but another not ? Does not this non-uniqueness itself demonstrate an algebraic inadequacy in the minimal rank theory ?

(*c*) *The Psychological Inadequacy of a Minimal Rank Solution.* The difficulties so far discussed are perhaps not insuperable. New and more precise methods might be developed for determining minimal rank and for selecting one set of communalities rather than another. An interesting recent development is the attempt to develop a minimal rank solution for the communalities without any prior estimate of rank. Tryon [16], Kaiser [9], and Guttman [6, 8] have all contributed to this line of approach. If their efforts are successful, minimal rank might be precisely determinable without the lengthy iterations used in this report; and that would indeed represent a considerable advance.

Their line of approach accepts Thurstone's formulation of the problem in terms of minimal rank (Kaiser and Guttman do so explicitly, and Tryon does so implicitly, as Guttman has shown). But the key question is surely whether the minimal rank solution so attained will be acceptable psychologically as well as algebraically. Present evidence suggests that minimal rank will generally be at least $[2p+1-\sqrt{(8p+1)}]/2$, and possibly more. And no good reason ever seems to have been provided for expecting empirical correlations obtained by sampling to fall into any simpler pattern than this. According to Ledermann's relation, at least 41 factors would have to be extracted when there are 50 tests, and 87 factors when there are 100 tests: that is to say, rigorous application of the minimal rank criterion, as proposed by the Tryon–Kaiser convergence procedure, is likely to supply far more factors than the psychologist would regard as reasonable. As a matter of fact, Thurstone, although the originator of the minimal rank theory, refused to go along with it when faced with its full implications. He writes, in relation to the Ledermann relation: "A battery of fifteen tests cannot be expected

to determine ten factors with any degree of stability; in order for a factor analysis to be stable and scientifically significant and convincing, the number of tests must be two or three times greater than those which are shown in the table" ([15], p. 294).

The mathematical difficulties of the minimal rank theory are well known; but less has been written about its psychological inadequacy. The results of this study—particularly the persistent Heywood cases—suggest that the mathematical model is not consonant with psychological requirements.

(*d*) *The Neglect of Statistical Considerations.* We thus seem forced to return to Thurstone's later suggestion that factoring cease when the residuals are small enough to be ignored: that is, the aim should be, not residuals of exactly zero, but residuals not significantly different from zero. In short the algebra cannot be divorced from the statistics.

The trouble with Thurstone's statement lies in its subjectivity. According to the way he presents it, each investigator is left to decide for himself when the residuals are small enough to be ignored. In this direction lies factorial anarchy. Fortunately, however, the sampling problem has been handled in the maximum likelihood method of Lawley [11] and Rao [13]. This seems to be the only method which meets statistical requirements and so supplies the objectivity lacking in Thurstone's own account. Rao explicitly refers to the minimal rank conception; and it would seem that there are two ways in which its central role in communality theory can be reduced. First, the requirement is now a minimal rank in the statistical not the strict algebraic sense. Rejected roots no longer have to be exactly zero; it suffices if they cannot be established to be significantly different from zero. Secondly, Rao is no longer concerned to find the loadings which fit correlations as closely as possible, but to find those loadings which correlate most highly with the observed test measurements: in short, the central emphasis is no longer placed upon the minimization of residuals.

Yet, even with the method of maximum likelihood, problems still remain. The procedure is difficult arithmetically even in these days of electronic computers. The communalities supplied seem to depend not only on the size of the sample of tests but also on that of the sample of persons. As more persons are added, more factors are likely to be established as significant, and communalities change. Unfortunately, as we have seen, the effect of additional factors is not simply to increase the various communalities: there is a restructuring; and some communalities decrease while others increase. The Lawley–Rao communalities therefore seem to be inherently tied up with the sampling. Perhaps, if some more serviceable method were found for determining the communalities associated with minimal rank, it might after all be shown that these form better population estimates of the communalities than the Lawley–Rao estimates, which are in fact biased on the lower side. But all this is a matter for further investigation.

The important point is that any solution of the communalities must be expressed in statistical rather than in algebraic terms. Thurstone's doctrine was inconsistent. The theory was formally started in terms of minimal rank;

and many of Thurstone's methods of communality estimation employ this framework. It was in these terms that the problem was expressed by Albert; and it is to this algebraic conceptualization that the investigations of Guttman, Heywood, and Ledermann have mainly related. However, once Thurstone had turned from artificial to empirical correlation matrices, the empiricist in him got the upper hand; and he himself re-wrote the problem in statistical terms. No doubt his own statistical exposition is incomplete, since no test of significance was provided: nonetheless it represented a far better statement of the communalities problem from the standpoint of the underlying logic.

References

[1] Albert, A. A. (1944). The minimum rank of a correlation matrix. *Proc. Nat. Acad. Sci.*, XXX, 144–148.

[2] Burt, C. (1952). Bipolar factors as a cause of cyclic overlap. *Brit. J. Psychol. Statist. Sect.*, V, 197–202.

[3] Cullis, C. E. (1913–1925). *Matrices and determinoids.* Cambridge, England : Cambridge University Press.

[4] Guttman, L. (1940). Multiple rectilinear prediction and the resolution into components. *Psychometrika*, V, 75–99.

[5] Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika*, XIX, 149–161.

[6] Guttman, L. (1956). Improved bounds for communalities. Research Report 6, 25 October.

[7] Guttman, L. (1956). To what extent can communalities reduce rank ? Research Report 8, 8 December.

[8] Guttman, L. (1957). Successive approximations for communalities. Research Report 12, 25 March.

[9] Kaiser, H. F. (1956). Solution for the communalities : a preliminary report. Research Report 5, 27 September.

[10] Kaiser, H. F. (1956). Communalities in factor analysis : a review. Research Report 10, 29 December.

[11] Lawley, D. N. (1940). The estimation of factor loadings by the method of maximum likelihood. *Proc. Roy. Soc. Edin.*, LX, 64–82.

[12] Ledermann, W. (1937). On the rank of the reduced correlation matrix in multiple factor analysis. *Psychometrika*, II, 85–93.

[13] Rao, C. R. (1955). Estimation and tests of significance in factor analysis. *Psychometrika*, XX, 93–111.

[14] Thomson, G. H. (1951). *The factorial analysis of human ability.* London : University of London Press. 5th ed.

[15] Thurstone, L. L. (1947). *Multiple-factor analysis.* Chicago : University of Chicago Press.

[16] Tryon, R. C. Communality of a variable : formulation by cluster analysis. *Psychometrika* (in press).

[17] Wilson, E. B., and Worcester, J. (1939). The resolution of six tests into three general factors. *Proc. Nat. Acad. Sci.*, XXV, 73–77.

[18] Wilson, E. B., and Worcester, J. (1939). Team tests for generals and specifics. *Proc. Nat. Acad. Sci.*, XXV, 358–364.

[19] Wrigley, C. F. An empirical comparison of various methods for estimating communalities. *Educ. psychol. Measmt.* (in press).

APPENDIX

A Note on the Correlation Matrix for Emotional Traits[1] : (*By C. Burt*)

The original object of the investigation referred to by Dr. Wrigley was to see how far the factorial procedures that had been introduced for the study of cognitive and educational abilities would also be applicable to assessments for motivational traits. Assessments were obtained for several groups both of adults and of children (*Brit. Ass. Rep.*, 1915, pp. 695 f.). One of the smaller correlation tables for children (comprising 8 variables only) has been factorized by Holzinger and Harman (*Factor Analysis*, p. 176) as well as by Dr. Wrigley, who extracted 8 factors for the 8 variables (this *Journal*, VII, p. 26).

The table reproduced by Dr. Wrigley in the foregoing article (p. 39 above and Table I of the paper cited, [2]) was obtained by averaging the coefficients from the four largest groups, amounting in all to 287 children. Let us first consider how far the values, reached by Dr. Wrigley in his remarkably thorough analysis, agree with those given in that paper ([2], p. 201, Table II). In the table below the first two lines show the communalities reached by Dr. Wrigley (Table VI, columns 2 and 3) and the last line those given in the original analysis. As will be seen, the agreement between the two sets of values is very close. Moreover, in several cases where one of my figures falls short of Dr. Wrigley's (e.g. that for variable 3 where my figure is 0·922 and his is 0·946), or again where mine is larger than his (e.g. that for variable 10 where my figure is 0·347 and his is only 0·338), one of his later procedures gives figures departing from mine by much the same amount *on the other side* (viz. 0·918 and 0·351 for variables 3 and 10 in Table V, column 3).

Table X. Communalities for Seven Factors

	1	2	3	4	5	6	7	8	9	10	11	Total
I	1·004	0·999	0·946	0·761	0·678	0·698	0·475	0·595	0·480	0·338	0·521	7·495
II	1·003	0·998	0·946	0·758	0·681	0·710	0·458	0·595	0·481	0·339	0·524	7·499
III	0·998	0·997	0·922	0·783	0·667	0·697	0·482	0·628	0·467	0·347	0·515	7·503

Note. Lines I and II give values obtained by Dr. Wrigley—line I with zeros as the starting-point, line II with squared multiple correlations. Line III gives the values obtained by Burt starting with rough trial-values inserted by inspection to fit the general pattern of the correlation table.

The method adopted in my own analysis was ' simple summation ' (based on the formula printed in *Distribution and Relations of Educational Abilities*, 1917, p. 53) ; and, as stated in the original paper, the final values for the communalities (and therefore for the factor saturations) were obtained by successive approximation : the details of the iterative method procedure are explained more fully in *Factors of the Mind*, Appendix I, pp. 449 f. : a desk machine was used for the computations. In deciding what figures to take as the trial estimates for the several communalities, the reliability coefficients (when available) were treated as marking the upper limit. But in practice the simplest procedure is to consider by inspection what figure will best fit the general pattern. For this purpose the variables are arranged of descending totals (as in my own table and in the table as reprinted above) : then, in the first few columns we insert figures somewhat *larger* than the largest correlation in each column, and in the last few columns figures somewhat *smaller* than the largest

[1] As the earlier publications giving the details of the researches on emotional traits may now be inaccessible to many overseas readers, the following note has been added at Dr. Wrigley's request.

correlation[1]. The factoring is usually continued until all the residuals are well below the level of significance.

In my view there is no one sovereign rule, applicable in every case, for determining how many factors should be extracted. With the present table only the first three factors were fully significant, and only three were given in the first report. If, however, we are intending to rotate a summational factor matrix containing (say) three factors only (one general and two bipolar) with a view to securing a group factor matrix, then we must be prepared to expect at least four group factors (possibly six if there is overlapping) in addition to the ' basic ' factor—i.e. five to seven factors in all. But this is not the final consideration. If we assume (as we must in this case) that each observed variable also includes a ' unique ' factor (i.e. a factor specific to each trait), then the principle of parsimony forbids us to extract *more* summational factors than would be required to reduce the correlation matrix to minimal rank with each communality < 1 (cf. Burt and Banks, ' The Reduced Correlation Matrix ', this *Journal*, VII, pp. 107 f.).

In determining the rank the assumptions on which I proceeded were as follows. " If the number of tests is a one of the ' triangular numbers ', 3, 6, 10, 15, . . . , the number of factors will be the next smallest triangular number, 1, 3, 6, 10, . . ." (loc. cit, p. 113). Hence with a 10×10 correlation matrix it would be possible to obtain a unique solution with six factors, and the number to be extracted should not exceed this figure, unless some of the communalities would then exceed unity. With an 11×11 correlation matrix the factor matrix for six factors would contain $11+10+9+8+7+6=51$ independent factor saturations. The correlation matrix, however, contains 55 observed coefficients. If we assume that each of these represents an independent value (which in my view is not necessarily the case), we should need a seventh factor. But in that case we should have to determine 56 unknowns with only 55 independent determinantal equations. Now, when calculating factor saturations for three summational factors only, the communality for the first variable (and to a less extent for the second) showed a marked tendency to exceed unity[2]. It appeared, however, that this could be readily counteracted by introducing supplementary factors with high weighting for the last variable (variable 11, for which the communality also showed a marked tendency to rise above its expected value, as is evident from the table). Accordingly, as was noted in my article, " the self-correlations [i.e. ' communalities '], reached by successive approximation, were computed on the principle that seven factors

[1] The communalities for the last few variables are always difficult to assess ; if the correlations in each of the last few columns descend regularly in size, the communalities will be smaller than the *smallest* correlation ; if, as here they rise towards the bottom, they will be larger than the nearby correlations. Except for rounding, the trial values guessed in advance by a fairly experienced factorist will, as a rule, approximate pretty closely to the values of the squared multiple correlations, but will in general be somewhat larger.

[2] It is interesting to speculate why in certain tables some of the communalities obtained with a reduced number of factors occasionally display a marked tendency to rise above unity. This peculiarity seems to occur most frequently in tables obtained by averaging coefficients from separate groups—a procedure that should be used only with the greatest caution. Yet it is by no means confined to such tables. The main reason, I take it, is that the number of operative factors must in actual fact always be very much higher than the number of significant factors : when ' unique ' factors are present, it must appreciably exceed the number of variables. Moreover, the distinction between ' common ' factors and ' unique ' factors is never very sharply drawn. ' Unique ' factors are usually regarded as including the ' error factors ' responsible for ' unreliability ' ; and it was shown long ago that the errors of observed correlations are themselves correlated (cf. K. Pearson and L. N. G. Filon, ' On the Probable Errors of Frequency Constants ', *Phil. Trans. Roy. Soc.*, CXCI, 1898, equation 37). Hence when ' unique ' factors are excluded from the communalities, their effects (the overlapping of errors) gets absorbed into the common factors, and thus tends to swell their size. This indeed is one of the arguments for occasionally working with tables of ' smoothed ' correlations instead of with those actually observed.

would be needed to account for a 11×11 table " ; and as the 56th datum I took a fixed communality of 0·998 for variable 1, to keep its value just below unity, and proceeded to compute the saturations for seven factors by successive approximation.

This particular correlation table was one of several used as the basis for practical exercises in factor analysis for successive generations of students ; and the communalities obtained by each student in turn were those adopted by his successor as the new starting point. With the values finally achieved it was found that a reasonably good fit was given by five factors : " the largest residuals then remaining are less than $\pm 0{\cdot}015$ " ; and, when the usual tests of significance were applied, it turned out that " the first three factors alone were fully significant ", as had been assumed in the first report.

May I add that I entirely agree with Dr. Wrigley in rejecting the principle of minimal rank as a criterion for the number of factors or for the most suitable communalities ? My own view was stated in the paper already cited : " Thurstone's statement that the reduced self-correlations or ' communalities ' must be so determined that the number of factors is equal to the rank of the side correlations appears erroneous : what we have to determine is not the smallest number of common factors that will maintain the rank, but the smallest number of *significant* common factors required to render the residuals non-significant. (In any case, unless the number of trials happens to be one of the triangular numbers, this principle fails to provide a unique solution). It would be better still to apply the principle of parsimony to the number, not of significant *factors*, but of significant factor *saturations*. If we assume that a set of summational, i.e. general-plus-bipolar, factors is sufficient, then the resulting number will, as a rule, remain unaltered. But, when we are not so restricted, a matrix of group-factors may give a more economical fit, for then the number of saturations may prove smaller even when the number of factors is not " (loc. cit., pp. 116–117).

Vol. XII Part I The British Journal of Statistical Psychology May 1959

EXPERIMENTS ON TELEPATHY IN CHILDREN

Critical Notice

By Cyril Burt
University College, London

The Mind Readers. By S. G. Soal and H. T. Bowden. London : Faber and Faber, 1959. Pp. 292. 30*s*.

The desire to secure more conclusive evidence for or against extra-sensory perception and allied phenomena has given rise to a series of researches based on statistical techniques. The experience so gained and the criticisms of the methods employed throw fresh light on the uses and limitations of similar techniques in other fields of psychological research, especially on the validity of the commoner methods of testing significance and on the results reached by the matching-method.

The new experiments reported by Dr. Soal in his latest book make a definite contribution to these problems, and yield prima facie evidence for telepathic communication far more striking than any hitherto obtained under conditions equally stringent. At the same time the observations incidentally collected do much to elucidate the factors involved.

The following paper examines the evidence thus adduced, and argues that the processes involved appear to be processes of thought rather than (as hitherto assumed) processes of perception. Alternative theories are briefly reviewed ; and suggestions put forward for future research, first in regard to hypotheses that now call for experimental testing, and secondly in regard to standardizing both the experimental and the statistical procedures used in such investigations.

I. Methods and Results

As every student of individual psychology is aware, whenever we wish to investigate some rare or abnormal phenomenon—mathematical prodigies, cases of word-blindness, or the resemblances of identical twins—the easiest plan is to start with children of school age. Children have more time to spare for the requisite experiments; their mental processes are simpler; and suitable cases can be most readily located by means of school surveys. It is therefore a matter for some surprise that almost all the work on paranormal capacities has hitherto been carried out with adults[1]. Partly no doubt this has resulted from a widespread

[1] A scheme for such a survey, in cooperation with a few interested teachers, was submitted to the London Education Authority as long ago as 1925, with support from the Chief Inspector, Dr. C. W. Kimmins. After prolonged discussion, however, it was decided that the project might lead to some misunderstanding ; and it was therefore abandoned. Quite recently, however, Dr. J. G. van Busschbach, Chief Inspector of Schools, Amsterdam, has carried out tests with pupils aged 10–12 as ' percipients ' and the class teacher as ' agent ' (the teacher being placed out of sight of the children at the back of the classroom). The number of correct guesses was significantly

notion that dabbling in such matters is itself the mark of an irrational and ill-balanced disposition, and apt to induce a similar attitude in the participants; but partly, I imagine, it is due to the fact that no direct evidence for such capacities among children has so far been forthcoming. For this reason Dr. Soal's new researches are of outstanding interest.

During a visit to Spain a few years ago, he and Mr. Bateman obtained modest but encouraging results with a peasant girl of fifteen named Maria[1]; and it thereupon occurred to him that unsophisticated children, living in remote rural communities, might prove far more likely subjects than inhabitants of large cities or students at universities on whom he had previously experimented with little or no success. He had himself, for over thirty years, spent many of his annual holidays with a family named Jones who lived in a small cottage in the Snowdon district. Two younger members of the family, cousins called Ieuan and Glyn, he had known almost from infancy. And on making a few informal trials, he quickly found that Glyn, the younger, displayed quite unusual powers when working as receiver or 'percipient' with his cousin as transmitter or 'agent'. At the time the experiments began the boys were 13 years of age with about 3 months between them. Both lived a rather busy life. Glyn attended a county school some distance away, and in his spare time earned a little extra money delivering papers to scattered farms. Ieuan left school while the experiments were in progress, and took a job as kitchen-boy in a hotel near by. Their I.Q.'s were approximately the same, 110. But they were of very different temperaments. Glyn was active, sturdy, and at times somewhat intractable; Ieuan quieter, more delicate, but more cooperative. Both were tested with the Rorschach test by a London psychologist who reported that, while Glyn was socially well adapted, Ieuan seemed by his scores a "neurotic introvert, out of touch with reality". Dr. Soal and his collaborator question the reliability of such verdicts; and, if (as I imagine) the Rorschach results were assessed in the ordinary way, I should readily agree. The investigators' own impressions appear far more trustworthy.

It would hardly be expected that the boys would willingly sacrifice their few free hours to the monotonous business of card guessing, without some strong incentive; and indeed they soon became restless and bored. They were therefore promised monetary prizes for every set of guesses well above chance expectation. The effect was so surprising that a long series of experiments under carefully controlled conditions was planned, and, in spite of numerous difficulties, successfully carried out. The whole research involved frequent journeys of 250

(though not greatly) above chance expectation. The experiments were repeated by Professor Tenhaeft at Utrecht and by Dr. van Busschbach himself in North Carolina with similar results : with older children, or with an agent who was not the children's teacher, no success was obtained (*J. Parapsychol.*, XX, 1956, pp. 131f, and *Tijdschrift voor Parapsychologie*, 1952–6 : cf. also M. Anderson and R. White, 'A Further Investigation of Teacher–Pupil Attitudes and Clairvoyance Test Results,' *J. Parapsychol.*, XXI, 1957, pp. 81–97).

[1] She was one of three Spanish servants employed by Mr. and Mrs. Langdon-Davies, who had already discovered that the girls apparently possessed telepathic gifts. Fuller details will be found in the paper by J. Langdon-Davies (*CIBA Foundation Symposium on Extrasensory Perception*, 1956, pp. 111–130).

miles either for the experimenters who lived in London or for the boys and their parents who were subsequently persuaded to make short visits to the metropolis. Part of the necessary expense was generously borne by the Parapsychology Foundation of New York.

In most of the experiments a pack of 25 coloured cards was used representing 5 different animals with 5 of each type in the pack. By chance therefore one would expect only 5 in each pack to be correctly guessed. In all over 600 runs were made. The results were spectacular. During the main period, when working at his best (August 1955 to April 1957), Glyn, the star performer in this new research, achieved 2 runs in which *every* card was correctly guessed; on 4 occasions he got 24 right out of 25; and he scored 20 or more correct hits on at least 26 occasions. Except for one or two claims unauthenticated by witnesses[1], there is nothing like it in the whole history of psychical research. Performances as remarkable as these must therefore be subjected to the closest scrutiny.

Experiments and results of this nature raise interesting statistical problems. At most of the sittings 8 runs with a pack of 25 cards were commonly planned, though not always completed. With $8 \times 25 = 200$ trials the mean chance expectation would be 40. If the number correctly guessed is (say) 64, then the crucial problem is to estimate the likelihood of obtaining by chance a deviation of $64 - 40 = +24$. For a run of 200 the usual mode of calculation yields a standard deviation of $\sqrt{(\frac{1}{5} \cdot \frac{1}{5} \cdot 200)} = 5{\cdot}66$. The 'critical ratio' is therefore $24 \div 5{\cdot}66 = 4{\cdot}2$; i.e. the deviation is over 4 times the standard error. From tables for the normal probability distribution we should infer that the odds against so high a deviation (positive or negative) occurring by chance would be about 50 000 to 1.

When dealing with sittings where high scores are *repeatedly* obtained Dr. Soal suggests an alternative procedure. Consider the series obtained on 3 January 1957. This was carried out in the open air with the boys separated by a distance of over 100 feet: (159 feet for the last two runs). Here the number of correct guesses in the 8 runs were 21, 19, 4, 19, 21, 7, 16, 20—a total of 127 hits out of 200—63·5 per cent instead of 20·0 per cent. To compute the precise odds would be extremely laborious and indeed scarcely necessary. However, suppose the probability of some specified result being obtained by chance during a *single* run is p where p is a very small fraction: for example, let the result be a score of 19 or more correct guesses during a single run of 25. Then, using

[1] By far the largest average score on record is that of an unnamed music teacher, brought to Professor B. F. Riess (of Hunter College, New York) by one of his students. After synchronizing watches, Dr. Riess turned up cards in his own house; and the 'percipient' wrote down her guesses in her room a quarter of a mile away: 37 double runs (1850 responses) were completed in this way, and the average number of correct guesses was 18·2. After a 'nervous breakdown', the 'percipient' resumed her tasks; but the number of correct guesses (average 5·2) fell to the chance level, and, at the end of 5 sittings, she refused to continue. Unfortunately no observers were present, either with the agent or with the participant; and Dr. Riess himself apologizes for "the many uncontrolled factors". With the rest of the well authenticated experiments the average number of correct guesses seems usually to have been between 7 and 10 (28–40 per cent).

the method described in the appendix (see p. 91), we should find that p was well under 10^{-8}, i.e. 1/100,000,000.

The probability that the event will happen r or more times in m trials will therefore be

$$\frac{r!(m-r)!}{m!}p^r$$

plus terms of a higher power in p: here r (number of scores of 19 or more) = 5, and $m = 8$. Hence the probability required is 56×10^{-40} plus terms in 10^{-48} (which may be safely ignored). Thus the *lower limit* for the odds against a chance hypothesis is the order of 10^{38} to 1. Perhaps the easiest way to indicate the magnitude of such a number is to point out that Eddington's estimate of the total number of electrons in the universe is of the order of $(30 \times 10^{38})^2$. Thus the odds are, quite literally, 'astronomical'. Nor was this particular run in any way unique. Sittings with correct scores amounting to over 60 per cent are recorded on several occasions. Two shorter sittings yielded scores of 19, 24, 20 and 19 (82 per cent) and 19, 20, 23 and 25 (87 per cent) respectively.

However, where a large number of sessions have been completed, probably the most convincing method, as well as one of the simplest, is to compare the names of the cards actually guessed, not with the names of the cards shown to the transmitter during the *same* run, but with those shown during another run, i.e. with an actual selection which the percipient was *not* trying to guess. Such a comparison yields an *empirical* estimate of the results likely to be obtained by chance. The printed scoring sheets on which the results were entered were designed to contain two runs of 25, the initial letters for the cards transmitted or guessed being entered in the appropriate column. Thus it was quite easy to match the cards guessed during the first run with those transmitted during the second run, and vice versa. The possibility that telepathy might be in some degree retrospective or anticipatory can be ruled out by comparing the results of the empirical matching with the theoretical value: (where such a phenomenon is suspected a more elaborate scheme of cross-checks could, of course, readily be carried out).

This procedure was adopted for all the indoor experiments undertaken between August and October 1955, and all the open air experiments undertaken between April 1956 and January 1957. They include practically all the trials carried out under satisfactorily controlled conditions. With the indoor experiments 5024 trials were completed. In theory we should therefore expect approximately 1005 successful hits; the cross-check with unrelated columns gave 998 cases of agreement. Actually the number of correct guesses amounted to 1889, 37·6 per cent instead of 20·0 per cent: this is 31 times the standard deviation. With the outdoor experiments the total number of trials was 4474; the theoretical expectation 895; cross-check 940; actual hits 1923, i.e. 43·0 per cent instead of 20·0 per cent, or 38 times the standard deviation. It would be quite incredible that these high proportions of successes should be due to chance.

II. Examination of the Arguments

Prima facie then these results seem to establish an extremely strong case for the hypothesis of 'extra-sensory' perception. Indeed those psychologists who advocate a 'hypothetico-deductive' procedure as the foundation of scientific methodology ought now to accept it as definitely proved; for the phenomena which are to be experimentally tested are logically deducible from that hypothesis and the experimental tests confirm the deductions. Nevertheless, the majority of scientists, I imagine, and a large proportion of present-day psychologists, will hesitate to accept Dr. Soal's conclusions as they stand without a closer scrutiny of the detailed evidence. The arguments commonly adduced by sceptical critics are of two main types: let us call them for brevity the *a priori* and the *a posteriori*. Taken together they imply that we have to assess, not a single posterior probability (namely, that the observations are due to chance coincidences), but several antecedent probabilities and several different likelihoods, and then combine them along the lines of Bayes' well known theorem. This is certainly the methodological standpoint I myself should adopt[1].

A. *The A Priori Arguments.* The *a priori* type of argument against extra-sensory perception has recently been stated in clear and forcible fashion by Mr. Hansel[2].

As a consequence of the existing body of scientific knowledge we are, he says, compelled to "assume that telepathy and clairvoyance are impossible"; we are therefore bound to seek some alternative explanation: and this, he argues, must be "some kind of trick, which involves only natural processes". He admits that the facts as recorded render this alternative improbable in itself; nevertheless he thinks that the antecedent presumption *against* a paranormal hypothesis is so overwhelming as to outweigh any *a posteriori* evidence.

But how can Mr. Hansel be so sure that telepathy is "impossible"? Admittedly no one can deny that cases do exist where absolute impossibilities can be affirmed in this way on purely *a priori* grounds. For example, in books on mathematical puzzles we are told how to construct what seems to the ordinary naïve observer a 'perfect' seven-sided solid; but a little mathematical training will suffice to yield a direct *a priori* insight into the fact that a regular heptahedron is indeed impossible at least in Euclidean space. Here therefore we can be certain that some trick is involved, even when we cannot actually detect it. But we have no such *a priori* insight into the fact that telepathy is impossible. And in dealing with any empirical issue it would I fancy, be generally conceded that we can never declare in advance that such and such a hypothesis is absolutely impossible, but only that it is highly improbable. Mr. Hansel, however,

[1] See C. Burt, 'Definition and Method in Scientific Psychology', this *Journal*, XI, pp. 60f.

[2] C. E. M. Hansel, 'Experiments on Telepathy', the *New Scientist*, V, 1959, pp. 457f. I select this paper because it is the most recent. With varying degrees of emphasis much the same arguments have been put forward by earlier writers. They are summarized with references in Tyrrell, *Science and Psychical Phenomena* (1938, Pt. iii).

endeavours to support his initial assumption by maintaining that the contradictory assumption would be logically incompatible with the four "basic requirements of science"[1]; and these basic conditions he apparently regards as self-evident axioms.

As he states them, the four conditions run in pairs. (1 *a*) All events are causally dependent on preceding events in the world of space and time, never on any future event, and (1 *b*) this causal dependence is based on the 'laws of mechanics', i.e. there must be a mechanical chain of causation, with no discontinuity in either space or time, connecting the preceding event or 'cause' with the dependent event or 'effect'. Of these two postulates the former excludes pre-cognition (of the type suggested, for example, by Soal's experiments with Shackleton), since that would imply that a future event determines knowledge of it before it has happened; the latter excludes clairvoyance and retro-cognition, since the one implies action at a distance without transmission of energy across the space separating stimulus from sense-organ, and the other implies action across a temporal interval, with no persistent structural effect in the brain to account for it. The other two conditions are more specific. (2 *a*) Conscious processes can directly determine, and are directly determined by, material processes in the brain or nervous system of the person who experiences them, and by nothing else; (2 *b*) the only way in which one person can convey information about his conscious experience to another is by material means, i.e. by moving his own material body (as in gesture, speech, or the like) and so transmitting energy to the receptor organs (eye, ear, or skin) on the surface of the other person's material body, and so to his brain. These further assumptions preclude not only psychokinesis, but also any form of telepathy, since (2 *a*) implies that it is impossible for B's mind to act on A's *brain*, and (2 *b*) implies that it is impossible for B's mind to act directly on A's *mind*.

As Mr. Hansel words them, the last two principles (2 *a* and 2 *b*) seem inconsistent with the first two (1 *a* and 1 *b*), for they recognize the existence (*a*) of 'mental' processes as distinct from physical or material processes and (*b*) of 'knowledge' which one person has of another person's conscious experience or 'thoughts' in the shape of "*information* reaching him through the senses" (Mr. Hansel's words but my italics). But 'mental' processes in general, and the acquirement of 'knowledge' or 'information' about another's 'thoughts' in particular, are not processes which obey the 'laws of mechanics'; and we

[1] Mr. Hansel states his four basic principles in the form of headings. But, since they are either postulates or empirical generalizations, they must be expressed in the form of propositions. The four he chooses seem to have been suggested by Professor Broad's list of the four 'basic principles' with which the 'established results of psychical research' are in ostensible conflict (C. D. Broad, 'The Relevance of Psychical Research to Philosophy', reprinted in *Religion, Philosophy and Psychical Research: Selected Essays*, 1953, pp. 7–26). Dr. Broad states them as 'limiting' principles (asserting what is impossible); for him therefore they are merely *necessary* conditions which our interpretations must obey if they are to fit in with the familiar mechanistic model; Mr. Hansel treats them as *sufficient* conditions. Moreover, he omits to point out that Dr. Broad's own conclusion is that, quite apart from other considerations, the "established results of psychical research . . . call for very radical changes in these basic principles".

could only regard them as events depending exclusively on events in the world of space and time, if the preceding events in that world are supposed to *generate* knowledge. And this means generating consciousness, for without consciousness there can be no knowledge. But a result such as this would imply the production of events without loss of energy. Hence to be consistent Mr. Hansel would have to adopt a thoroughgoing materialistic standpoint, like that of Watson, expunging such words as ' mental ', ' consciousness ', ' knowledge ', and ' thoughts' from his vocabulary, and referring solely to ' behaviour ', i.e. to bodily movements in the world of space and time. Of course the facts of so-called extra-sensory perception could quite easily be described in purely behaviouristic terms; yet even then they would still conflict with principles (1 *a*) and (1 *b*). That indeed is why they are termed ' paranormal '.

Now Mr. Hansel's interpretation of what he regards as the basic principles of science would have been perfectly acceptable in the heyday of the great Victorian physicists, when it was assumed that the universe could be explained completely in terms of matter, energy, and an ubiquitous ether which acted as a conveyor of energy, that every event was caused by some preceding event continuous with it in a world of absolute time and space, and that all causal laws (even those of electricity, magnetism, and gravitation) could eventually be reduced to the ' laws of mechanics '. But this simple cosmology is as obsolete as the bonnets and bustles of Queen Victoria herself. In these days of relativity and quantum theory we can no longer regard the old ' basic requirements ' as self-evident axioms to which the whole universe must conform, as Mr. Hansel would have us do.

Moreover, the ' laws of mechanics ' to which he refers[1] are merely approximate laws, accurate enough when dealing with ' mesoscopic ' objects like the earth,

[1] I find it difficult to understand exactly what Mr. Hansel means by his phrase ' the laws of mechanics '. His reference to ' causes ', and his emphasis on the way all events must be dependent on events preceding them in time, suggests that he is referring to the ' laws of motion ' as formulated by Newton and his successors (a sense which the phrase very frequently bears). Whether he would also include Newton's law of gravitation (which, like telepathy, involves *prima facie* a ' force ' or cause acting at a distance) is not altogether clear. The ' laws ' (or equations) of motion assume that the behaviour of two or more material bodies (i.e. particles or aggregates of particles) depends solely on (i) the spatial position of each body relative to the others, (ii) the relative velocity of the bodies, and (iii) the time interval. The differential equations are then of the second order. The telepathic phenomena described by Dr. Soal do not obey such laws, and violate each of these three assumptions. If we use Lagrange's equations, we substitute energy for quasi-causal forces, and obtain equations of the first order, invariant as regards transformation of the coordinates. By incorporating Hamilton's principle (a generalization of the principle of least ' action ') we secure an alternative formulation of still wider generality. Moreover, as has often been remarked, ' the Newtonian formulation ' (which is expressed by *differential* equations) " stresses the causality aspect of nature ", the more modern Hamiltonian formulation (which is expressed by an *integral* equation) " emphasizes that nature for ever tries to achieve a certain purpose or end, so that the resulting motion realizes this event, which is an event in the future " (R. B. Lindsay, *Concepts and Methods of Theoretical Physics*, 1951, p. 136). Planck's quantum theory (1900) was in flagrant contradiction with Newtonian mechanics, but down to about 1923 could be expressed in terms of Hamiltonian mechanics; and it is instructive to recall that Planck, to the end of his life, looked on the validity of integral principles as " monumental demonstrations of purpose in the universe ", and so superseding the simple type of deterministic cosmology which Mr. Hansel still postulates.

the sun, and the various planets (except Mercury), and with man-made machines and projectiles. They are working rules for a familiar world whose lower limit is the indivisible atom and whose upper limit is the edge of the Milky Way. They break down completely when applied to bodies outside this familiar range of size or moving with unfamiliar velocities, e.g. electrons, galaxies remote from our own, and, we must surely add, the intra-molecular processes of the human brain. Even as late as 1920 there were still a few physicists who hoped that the 'laws of mechanics' might yield a complete theory of the atom. But no one would subscribe to such a view today. And, although many psychologists still seem, like Mr. Hansel, to imagine that co-called paranormal phenomena would be "in conflict with physical science", hardly any contemporary physicists familiar with the facts would endorse such a pronouncement[1]. Physical science itself is in too unstable a state for anyone to reject out of hand an internal coherent hypothesis *merely* because of an alleged conflict with physical science. But in point of fact the conflict to which Mr. Hansel alludes is not with physical science but with materialistic science[2].

The aim of physical science is to deduce as wide a variety of facts as is possible from the smallest number of the simplest postulates. Its postulates therefore are laws of a theoretical model, not laws of nature. Mr. Hansel would be fully justified if he treated his own 'laws' as stating what would be impossible in the particular theoretical model he proposes to try out first of all; but he is certainly not justified as treating them as infallible axioms incapable of modification or revision. Since "the physical processes in the nervous system" to which he refers are now known to depend in part on processes of a sub-atomic order, we can no longer insist that the 'mental processes' associated with them must necessarily conform with a version of the 'laws of mechanics' which no longer holds good within the atom. Accordingly, we are entirely free to construct hypotheses that violate every one of Dr. Hansel's four principles, if the facts require it. And indeed, as we have seen, most of them are already violated by accepted hypotheses in other fields.

However these newer hypotheses have only been adopted because simpler alternatives have been disproved by a long series of experiments, and because they now enable us to coordinate, and sometimes to predict, a large mass of phenomena in ways which otherwise would be impossible. Nor should I for one moment deny that any hypothesis which tends to increase either the number or the complexity of the postulates involved has a fairly high *a priori* probability against it. Whether in the end it can be accepted must therefore depend on the posterior probability in its favour as compared with that of the alternative hypotheses. Accordingly, in considering how to account for such phenomena as

[1] Curiously enough it was the leading physicists—Crookes, Barrett, Lodge and others—who were the first to accept the findings of psychical research at their face value. It is rather from biologists and physiologists that opposition has chiefly come.

[2] Unfortunately writers on psychical research have been guilty of the same confusion and have themselves been only too anxious to insist that " paranormal phenomena are incompatible with the doctrines of physical science" (cf. Tyrrell, *loc. cit. sup.*, p. 353).

those now reported by Dr. Soal and Mr. Bowden our next task will be to examine the probability of the various alternatives.

B. *The Alternative Hypotheses.* There are at least half a dozen conceivable explanations that require to be examined.

1. Let us consider the arguments of the statistical critics first of all. In several publications Mr. Spencer Brown has argued that ESP phenomena are simply statistical artefacts, due to a defect, hitherto ignored by practical investigators, in the accepted theory of probability and chance. His main contention is that the concept of randomness is purely theoretical, and applies only to an infinite series of events; since therefore all actual series are necessarily finite, we can never assume that they conform to a strictly random distribution. He has himself paired lists of ' random digits ', taken at haphazard from published tables, and in several of his experiments found that the number of coincidences exceeded the expected value by over three times the standard deviation[1]. Cruder methods of randomization may therefore yield a yet larger number of accidental matches.

Now these arguments are by no means as new as is commonly supposed; but they serve as a timely reminder that the psychologist must scrupulously observe certain necessary precautions in planning a statistical inquiry. (i) When a delicate test of significance is required, ordinary methods of randomization, e.g. card-shuffling, are apt to be unreliable. Kendall, for example, gives a table of the ' distribution of first hands at whist ' grouped according to the number of cards belonging to a given suit. A chi-squared test shows that the odds against the distribution being genuinely random are about 10 000 to 1. The explanation is that in actual play certain kinds of card may collect together in a way which ordinary shuffling does not sufficiently correct (*Advanced Theory of Statistics*, 1943, I, pp. 298–300). (ii) In every actual research some kind of selection is generally essential. Hence a list of random numbers intended for practical use should, so far as can be managed, give random results whether used in part or as a whole. But, as Kendall points out, "it is impossible to construct a table which will satisfy this requirement entirely" (*loc. cit.*, p. 196). Any actual series of numbers (e.g. a block of 25 digits all the same) may form part of a random series in the sense that it might arise from random sampling if the series were long enough. Accordingly Kendall and Babington Smith have indicated what parts of their table should be avoided in making comparatively small selections. (iii) As the history of scientific research warns us, almost every investigator is tempted to continue his experiments until he happens to get some

[1] G. S. Brown, ' Statistical Significance in Psychical Research ', *Nature*, CLXXII, 1953, pp. 154–6; cf. *id.*, *Probability and Scientific Inference* (1937). It is perhaps regrettable that, in spite of many requests, Mr. Brown has never published the full details of his ' mock ESP experiments ' and his ' random number counts '. Several empirical tests of his contentions have been carried out; and the general conclusion is that, except in borderline cases, the effects are not likely to be serious: (A. T. Oram, 'An Experiment into Random Numbers ' *J. Soc. Psych. Res.*, XXXVII, 1954, pp. 169; cf. J. A. Greenwood and C. E. Stuart, ' Mathematical Techniques used in E.S.P. Research ', *J. Parapsychol.*, I, 1937, pp. 206–225).

fairly definite result, and then very frequently stops. What is worse authors and editors are more likely to publish experiments yielding positive results than those that are purely negative. And, when this kind of policy is adopted, tests of statistical significance lose their value. It is the familiar fallacy of the visitor to Monte Carlo who prolongs his gamble until he has made a little money, stops before his luck has changed, and then comes back with the story that he has evolved a successful 'system'. The moral is twofold: first, the investigator should decide on the total number of trials he intends to carry out before he starts his experiments; secondly, he should also indicate, however briefly, the number of trials performed which proved abortive[1].

We should therefore be foolish to dismiss out of hand the criticisms Mr. Spencer Brown has advanced. Nevertheless they do not, as several critics of statistical psychology have argued, require statisticians to re-write their textbooks or psychologists to abandon such statistical methods and formulae as have been sanctioned by high authorities like R. A. Fisher for use in ESP investigations. At most they may perhaps shake our faith in conclusions drawn from data which deviate but slightly from chance expectation; and, it must be owned, most of the deviations hitherto obtained in paranormal research have been so inconspicuous that, apart from statistical tests, few observers would deem them worthy of notice. But the results now reported by Dr. Soal and his co-workers are so impressive that no one, however sceptical, could seriously think of ascribing them to the freaks of chance; and, if any doubt still remained, Dr. Soal's use of empirical estimates in his tests of significance would dispose once and for all of any suggestion that he trusts too readily to 'academic formulae'[2]. Accordingly, since the theoretical basis of his research seems sound enough, we must turn to the practical aspects, and ask how far we can rely on the care and accuracy with which the data were obtained and the figures analysed.

2. Those who know nothing of Dr. Soal or his previous inquiries might perhaps be tempted to wonder whether he may not, in his enthusiasm, have been led—perhaps quite unconsciously—to introduce a bias or error into his arithmetical and experimental work. Dr. George Price, an American scientist, has flatly declared that psychologists and others who publish such "incredible results are always to be suspected of collusion". And in a recent article discussing Dr. Soal's earlier investigations, Mr. Hansel has contended that, since the ostensible results are in conflict with the laws of physical science, they must

[1] Formulae which may enable us to allow for the effects of 'optional stopping' have been proposed by T. N. E. Greville ('A Summary of Mathematical Advances bearing on ESP Research,' *J. Parapsychol.*, III, 1939, pp. 86–92) and by J. A. Greenwood (*ap.* G. G. Pratt *et al.*, *Extrasensory Perception after Sixty Years*, 1940).

[2] Mr. Brown hardly does full justice to Dr. Soal's earlier work. For example, in the experiments with Basil Shackleton 'telepathic' conditions alternated with conditions for 'clairvoyance' without the percipient being aware of the difference: with the former a number of precognitive hits were obtained—the odds against chance being at least ten million to one; with the latter the scores were almost exactly what chance would suggest. How is it that the 'statistical artefacts' occurred solely under telepathic conditions? (Cf. S. G. Soal and F. Bateman, *Modern Experiments in Telepathy*, 1954, pp. 150–3.)

be due to 'trickery'[1]. Let me therefore say at once that Dr. Soal is a man of high academic reputation and an extremely competent mathematician: he was in fact formerly Senior Lecturer in Mathematics at Queen Mary College in the University of London. He is widely recognized as a critical investigator of alleged paranormal phenomena, and was elected to an appointment in Psychical Research at Trinity College, Cambridge. As many of his earlier experiments were carried out at University College, London, I myself had frequent opportunities of observing his work, and can vouch for his skill and scientific caution (which Dr. Rhine indeed has even characterized as 'overly defensive'). His own publications show how ready he is to acknowledge the frequent failures he himself has encountered after long, arduous, and expensive experiments.

In the present investigations at least three persons would have had to be parties to a very prolonged conspiracy, if there was one. Furthermore, the names of more than 25 persons are given who (in addition to parents and other relatives) watched many of the experiments. These include psychologists like Professor Mace and Dr. Thouless, mathematicians like Mr. Gliddon and Mr. Moakes, experienced investigators of psychical phenomena like Mrs. Goldney and Dr. West, scientists like Dr. Parkes F.R.S. and Dr. Perry of the National Institute of Medical Research, engineers expert in acoustics and radio, several schoolmasters, and a professional conjuror skilled in faked telepathy. They were allowed to choose the packs, shuffle the cards, supervise the records, and often to dictate the precise conditions. Signed statements are included from five of the best known of these observers, written in many cases immediately the sitting was over and testifying to the high scores observed and the scientific rigour with which the experimental conditions were controlled.

3. Probably, however, the majority of critics will be more inclined to question the reliability of the boys than that of the investigators. One difficulty was encountered at the very outset. Any attempt to place each lad alone with an adult supervisor in a separate room, behind closed doors, aroused clear signs of resentment[2]; concentration became difficult; and scores soon fell to their chance values or even below. On the other hand, so long as they remained within sight and hearing of each other, there were manifest opportunities for cheating.

[1] *The New Scientist*, V, 1959, pp. 457–460. The particular trick that Mr. Hansel suggests involves active collusion between the percipient and either the 'agent' or the experimenter. But, since in the experiments he cites, a number of different 'agents' took part, it would seem that he believes that Dr. Soal himself must also have been (as he puts) engaged in 'pulling our legs'.

[2] It is a tendency that has been noted not only by experimentalists who have tried to work with two or more school friends simultaneously, but also by those who have studied group-formation among children. Such a pair are very apt to form an 'in-group' of two, especially where an adult provides a foil. Country lads in particular are apt to get embarrassed or anxious when an investigator tries to separate them in the way described.

At this stage of the work, the experiments were being conducted in a tiny cottage, where only three rooms were available for the boys, the investigators, and the rest of the family; and, as the somewhat stormy scenes reported from time to time sufficiently show, the Jones family shared the common Welshman's feeling that the Saxon visitor may be seeking to exploit him, so that there is no reason why he in turn should not occasionally exploit the Saxon.

Nor indeed, as might be expected of lads of their age, were they above trying to exploit the generosity of a kindly if rather tiresome visitor, who imposed such tedious tasks and offered such tempting prize-money. It would seem that, after talking over these queer requirements both with each other and other Welsh youngsters, they concocted a kind of would-be 'secret code', which was, as might be expected with not very intelligent youngsters, exceedingly naïve. One day, when things were not going too well, they started to try it out. The trick was, of course, immediately detected. Nevertheless, the episode, occurring as it did at an early stage, was in many ways fortunate. First of all, had nothing of the sort been reported, those who know the average teen-ager's mentality would have felt a little uneasy: surely, it would be said, such boys were almost bound to try it on. But secondly, it evidently brought home to the investigators the supreme necessity of securing the most conclusive proofs that no such dodges were used again during the remainder of the experiments.

In theory, if we are to prove that some form of 'extrasensory perception' is at work, it is necessary to exclude every known means by which energy or information could possibly be transferred. Let us therefore briefly consider the chief alternatives.

(i) The use of a *visual* code was, so far as possible, excluded by shielding the transmitter's hands and legs from the view of the percipient. In the later trials an opaque curtain was erected, which completely hid each boy from the other. Throughout the experiments their faces and general behaviour were closely observed.

(ii) These devices would not preclude the use of some means of *auditory* communication. Several methods are conceivable.

(*a*) The simplest of all would be *whispering*. Since Ieuan (the transmitter) was closely watched by the investigator who gave the signals and sat close to the boy, any words would have to be formed without movements of the lips, and any message so faint as to be inaudible to the average adult. The skilled ventriloquist can of course produce articulate sounds with the lips slightly open, but apparently immobile; yet in such cases, close inspection (as with a near shot on the television screen) nearly always reveals some small degree of movement; and most performers therefore rely on a misdirection of attention. Dr. Soal, however, has himself clearly demonstrated that semi-articulate vowel-sounds, recognisable as words, can be whispered even with the lips completely still and closed. Moreover, such articulation need be neither deliberate nor conscious. As long ago as 1895 two Danish psychologists, by using two concave sound-mirrors, were able to prove that many persons, when thinking of certain numbers, involuntarily produced recognizable sounds in the nose or throat[1]. And it is a well-known fact that children have a far keener sense of hearing than adults,

[1] F. C. C. Hansen and A. Lehmann, 'Uber unwillkürliches Flüstern', *Phil. Stud.*, XVII, 1895, pp. 471–530. The results were later confirmed by J. L. Kennedy, *Psychol. Bull.*, XXXV, 1938, pp. 526f.

and inhabitants of the country a keener sense than those brought up in towns[1].

To check this possibility, three open air sessions were carried out in the playing field of St. Paul's School, during a visit of the boys to London. Here a large canvas screen was erected to shield each boy from the other's view; and the distance between the boys varied from 69 to 99 feet. Remarkable successes were still obtained. Moreover, if sound or any other physical method of communication were adopted, we should expect the amount of success to diminish with increasing distance. Actually for the shorter ranges the average score was 8·8 hits out of 25; for the longer 12·0. Both averages, it will be noted, are well above chance expectation.

(*b*) A possibility which has not been sufficiently examined in previous inquiries is the use of a *supersonic* code. Under laboratory conditions, where background noise is reduced to a minimum, many children can hear high-pitched sounds of 20 000 cycles per second or more even when the intensity is as low as 20 decibels; the average middle-aged adult can barely hear sounds of 8000 cycles at 40 decibels. (Normal speech involves sounds for the most part below 4000 cycles.) It might therefore be imagined that Ieuan could have used a code based on hissing or 'tongue-whistling' so faint and high pitched that no adult could perceive it.

Dr. Soal accordingly arranged for Mr. Alec Reeves (an expert in acoustics and Chief of the Research Laboratories of Telecommunications Ltd.) to sit within about 2 feet of the boy during several of the sessions. In a written report Mr. Reeves states that the hypothesis of auditory signals is completely negatived by the high scores made at 160 feet, and that no secret microphone could be manufactured which even theoretically would have enabled Glyn to pick up the assumed whispers at such distances[2].

But "the best counter to this supersonic assumption is that no tendency has been observed for the inverse square law to be fulfilled", i.e. there was no falling-off in the boy's successes with increasing distance.

(*c*) A code, however, might easily be devised based, not on sounds produced vocally, but on *noises*—taps with the foot or chair, coughs, sniffs, or the like. This was in fact the method adopted by the boys when they began to cheat. It was, as we have seen, immediately detected; and for the future the rigorous conditions just described entirely ruled out any repetition of the device even in a subtler form. On one occasion Dr. Soal suddenly changed to fresh cards, bearing

[1] The most striking example on record (if it can be taken at its face value) is that of the backward Latvian peasant girl who was able to 'read' an unknown text in a foreign language, when it was read 'silently' by her mother standing beside her: (cf. F. von Neureiter, *Wissen um fremdes Wissen*, Gotha: Klotz-Verlag, 1935, and Hans Bender, 'The Case of Ilga K.: Report on a Phenomenon of Unusual Perception', *J. Parapsychol.*, II, 1938, pp. 5–22). Cf. S. S. Stevens and H. Davis, *Hearing: Its Psychology and Physiology* (1938) and C. C. Bunch, 'Age Variations in Auditory Acuity', *Arch Otolaryng.*, IX, 1929, pp. 625–636.

[2] This would not apply to the supersonic whistles sometimes used by Welsh shepherds at sheep-dog trials. But such instruments could not be successfully concealed either about the person or during actual use.

symbols the boys had never seen before. The boys, without any comment, at once produced scores that were quite as successful as before; here therefore no code could possibly have been prepared or used. On another occasion 8 runs were made without Ieuan seeing the face of the card; all Glyn's scores were still on the same high level as before (average 9·6 instead of 5). He also succeeded with a so-called 'down-through clairvoyance' test (scoring 10 instead of the expected 5)[1]. Had they been relying on some concerted trick, their scores would have almost certainly dropped under these new conditions.

(iii) Music hall performers sometimes use *midget radio sets*; and both Professor Stratton, F.R.S. and Dr. Dingwall have independently put forward this explanation. Glyn, it was suggested, could have had in his pocket or shoe a few inches of wire which would serve as an aerial and produce felt vibrations rather than audible sounds. However, the smallest known transmitter, working with transistors instead of valves, would (with the requisite battery) have to be at least as large as a marble or cigarette holder; and, on strange premises like Birkbeck College or the playing field, Ieuan or the signaller would have to secrete this somewhere on his person.

To rule out this possibility the boys and Mr. Bowden changed into bathing slips and removed their wrist watches. Both they and the chairs and table were carefully searched; and, as it was summer, the runs were carried out in the open air at a distance of 166 feet. The average score was 63·3 per cent (as compared with the theoretical expectation of 20 per cent); on four of the runs the scores reached 20 or over. This surely disposes of the popular suggestion: 'It's all done by wireless'.

(iv) Those who readily acknowledge the moral integrity and scientific competence of the investigators and the various witnesses may nevertheless be tempted to suggest that, as mere scientists and statisticians, they could hardly claim the skill in detecting tricksters that would be possessed by a conjuror who has specialized in this field. In point of fact, Dr. Soal, Dr. West and Mrs. Goldney have probably had as much experience as anyone in exposing and studying telepathic fakes. Nevertheless, as Professor Stratton observed to Dr. Soal, the testimony of someone like Mr. Jack Salvin would be worth more than that of half a dozen scientists even if they were Fellows of the Royal Society like himself. Mr. Salvin is Chairman of the Occult Committee of the Magic Circle—a committee whose function it is to test the genuineness of claims to clairvoyance, mediumship, and telepathic powers, and to invent novel devices for the use of professional conjurors. He has studied the subject for over twenty years, has a first-hand knowledge of the codes and tricks of the most skilful illusionists, such as the Zancigs, the Piddingtons, and Marion, and is an expert

[1] In this the percipient is required to name in order the symbols on a pile of 25 shuffled cards which are untouched and unexamined until the test is completed. A few later runs under clairvoyance conditions failed, apparently because Glyn lost confidence when he realized the conditions had been changed; and Dr. Soal accordingly decided to concentrate on telepathic conditions.

in producing faked performances himself[1]. His own attitude towards telepathic phenomena has been extremely sceptical. Very generously he agreed to investigate the boys if they could be brought to London. On this occasion Dr. Soal was absent; and the arrangements were handed over to Mrs. Goldney, who impressed on Mr. Salvin that he was entirely free to impose whatever conditions he wished. Six sittings were carried out under his instructions. At the first three, scores of 43 per cent, 51 per cent, and 78 per cent were secured (including two scores of 20 or over). During the two concluding sessions, however, the boys (as usual after long spells) seemed to be getting bored; possibly they wanted to see a little more of London; anyhow the successes declined. But even on pooling the records of all the trials, successful and unsuccessful alike, the net result was amazing. Mr. Salvin in a signed report declares: "I am completely satisfied, after making all the observations I desired and having permission to do whatever I wished, that no trickery took place on the part of the boys or anybody else; in fact, under the experimental conditions I witnessed, trickery was impossible".

To baffle a group of three expert investigators of pseudo-telepathic phenomena for such lengthy spells—a feat which highly skilled professionals like the Piddingtons and the Zancigs have been unable to accomplish—is surely an astonishing performance for a couple of village boys. In view of all this evidence we must surely admit that every conceivable hypothesis based on ordinary methods of communication is completely excluded, and that the net probabilities are strongly in favour of some less familiar mental process. For convenience therefore let us refer to it as 'telepathy' with the proviso that the name in itself has no explanatory significance.

III. Conditions for Success

With all ESP performers one of the most constant features is the irregularity of the results: even with the best subjects one can never *rely* on getting a successful run. To this Glyn and Ieuan formed no exception. But with them, as we have seen, the long series of investigations was carried out under the most varied conditions; and a good deal of supplementary information could be obtained from the boys' parents and by other means, in a way hardly possible with adults. As a result considerable light is thrown on the factors making for failure or success. Above all, the experiments bring out very clearly what has often been suspected, though not so easily perceived in the case of older performers, namely, that the exercise of telepathic capacities is intimately dependent on the mental state of the performers.

[1] The wide variety of complex and semi-scientific devices used by present day performers are the product not only of many years of training (usually with the same partner) but also of the professional work of their predecessors. Mr. and Mrs. Salvin's act, which is carried out in the face of the most stringent conditions, gained a prize at the Paris Magical Congress. I believe I am right in saying that their method or methods have never been detected or divulged.

(i) *Concentration* was essential on the part of both boys. With Ieuan it seemed to turn rather on the inhibition of distracting ideas and emotions than on an intense voluntary attention to the card that was shown. His attitude was tense; but his face appeared dreamy and abstracted, as though his mind was empty of all thoughts except the picture before him, which, oddly enough, he often failed to recall a moment later. Glyn, on the other hand, when working well, showed more positive signs of concentrated thought; and unlike many adult percipients (Basil Shackleton or Mrs. Stewart, for example), he would commonly pause and consider with every appearance of effort.

(ii) That being so, it is not surprising to find that, with both, the ability was subject to *fatigue*. However, when they began to flag, a break of a day or two would frequently restore them. Moreover, with most of the sittings a little warming up seemed to be required before the successes appeared.

(iii) A favourable emotional *mood*, particularly in the case of Glyn, was indispensable. When he became sulky or resentful over some imaginary grievance, then, instead of waiting intently for impressions, he would call out the first thing that came into his head. With both boys the presence of an uncongenial stranger or any circumstance that tended to arouse fear or anxiety—being isolated in separate rooms, for instance, or wearing apparatus for electro-encephalographic records (which the parents seemed to regard as a possible instrument of torture)[1]—turned the experiment into a failure; and it was much the same if they were required to give up some more attractive occupation. In moods such as these the scores were often well *below* chance-expectation, as though they were deliberately avoiding the correct answers. On the other hand, an enjoyable outing or a walk over the hills was commonly followed by a successful sitting. The fact that the boys had known Dr. Soal for many years, and so accepted him as a long-standing friend, was manifestly one of the main factors. Glyn was tried with eleven other 'agents'; but the result was nearly always a failure except for a brief run with Ieuan's fourteen-year-old sister. Apparently success is obtainable only with a sympathetic agent, i.e. one who (from previous associations) inspires in the percipient the requisite feelings of comfort and self confidence.

(iv) The most striking feature was the supreme importance of a strong *motivation*. Unless both were keen to do well, the session was wasted. With Glyn vanity evidently played a part; he was especially eager to complete a whole run of 25 without a single mistake; and, as we have seen, eventually succeeded on two occasions. But with both boys the prize money, particularly when they had some rather expensive project on hand, seemed the most powerful incentive of all.

(v) When Glyn had failed under ordinary conditions, he still made high scores when under *hypnosis*. Apparently its chief effect was to inhibit boredom

[1] A few runs were carried out with the machine at the Belmont Mental Hospital; but no successful scores were obtained, and the boys' E.E.G. records showed no exceptional features.

and other unfavourable emotional conditions. Moreover, under hypnosis he could sometimes succeed with other agents besides Ieuan, and on one occasion even Ieuan became a successful percipient. However, contrary to what usually happened in the waking state, scores made under hypnosis began well, but rapidly declined as the boy dropped into a deeper sleep[1]. Post-hypnotic suggestion was effective on at least one occasion, but seemed highly unreliable. But so far only a beginning has been made with this particular procedure.

(vi) On the cognitive side the most important condition appears a *limitation in the range of objects* to be guessed. With free material Glyn failed completely; and himself remarked, "This wouldn't be possible: it might be anything in the world". The same restriction holds of nearly all successful card guessers, in spite of their wider claims: they can only say *which* object out of a limited number the agent is thinking of, not *what* object out of any large number[2].

Though poor in visual imagery, both Glyn and Ieuan appeared to belong to the intuitive or 'synthetic' type rather than the abstract or 'analytic'. Such differences may possibly be associated with what used to be described as 'differences of racial temperament.' Older writers were fond of emphasizing "the sensibility, the turn for natural magic possessed by the Celt" (to borrow Matthew Arnold's favourite phrases). Physically the inhabitants of the Snowdon district show distinctive anthropological characteristics; and this is traditionally an area in which paranormal gifts are supposed to be widespread. But far more important to my mind is the presence of a local climate of opinion favourable to such accomplishments. The abstract or analytic type of mind acquired by the student or the scientist tends to be highly sceptical of them; and scepticism in the performer himself, and any feeling that his own acquaintances would be somewhat contemptuous of such efforts, are always inimical to success. Readers familiar with *Aylwin* and its descriptions of the gifts possessed by the 'Cymric child' of Snowdonia, and still more those who have lived in the less frequented corners of that region and can gossip in the native tongue, will realize that there local opinion still takes such phenomena as a matter of course; neither of the boys, we are told, "thought there was anything mysterious or uncanny about

[1] I may perhaps observe that many years ago, when I myself was taking part in some hypnotic experiments at Oxford with Dr. McDougall, I succeeded in obtaining high scores with two young brothers, whom I knew quite intimately, when both were simultaneously hypnotized. I too found great difficulty in keeping both at the requisite light stage. As I turned from one to the other, the former would drop off to sleep, and vice versa. I have since had a similar success with two pairs of identical twins (girls); but their scores rapidly declined both during the individual sessions and when the series was repeated. From my own experience and from the comments of my subjects I got the impression that the effect of hypnosis was to allow the agent to concentrate on the card or number to be transmitted to the exclusion of all else, and to render the mind of the percipient a complete blank except for impressions received from the agent, with whom, through the hypnotizer, he was for the time being in close rapport.

[2] My own results suggest that the range of alternatives increases slightly with the intelligence of the percipient, perhaps owing to increased span of attention, and that it is desirable that both agent and percipient should be of the same type as regards the imagery used.

the business: they were just rather more 'lucky' than other people''[1]. On these grounds I should agree with Dr. Soal that his experiences fully support the suggestion that ''telepathy flourishes best among primitive peoples who live unsophisticated lives in close contact with Nature and tends to disappear in over-civilized and over-intellectualized communities''.

In certain respects, as Dr. Soal points out, the performances of the two boys ''refute several currently accepted dogmas''. Except where obvious motives were at work their powers showed no 'decline effects' nor did they yield evidence of 'displacement'. On the other hand, he says, ''they still leave us very ignorant in regard to the determining factors: extra-sensory perception still remains incomprehensible and unpredictable''. It is not in his view a specifically 'psychic' activity; and he can only conclude that for any real light on the actual processes involved we shall have first to learn more about the relevant facts of ''biology and brain functioning''.

Finally, it must, I think, be owned by every impartial reader that, alike for their success and for the care with which they have been conducted, the experiments here recorded are unrivalled in the whole corpus of psychical research; it is clear that here, as in so many other fields of scientific study, the one thing the investigator needs is an inexhaustible fund of patience, perseverance, and resourceful ingenuity to overcome the lurking obstacles that would quickly daunt most ordinary observers.

This bald synopsis of the main results obtained by Dr. Soal in these new investigations fails to do justice to his own account. He relates his experiences, not in the bleak schematic form I have adopted here, but as a chronological narrative having much of the suspense and surprise of a good detective story and containing vivid character-sketches of the personalities involved. His account breaks off rather abruptly; and for the present, it would seem, he himself is keeping an open mind as to ultimate explanations, until he has continued his experiments either with the same two subjects or with others who may prove equally promising.

Were it not for the strong antecedent presumption *against* extra-sensory methods of obtaining factual information, his results would, I imagine, be hailed as virtually conclusive. As it is, an impartial critic is bound to insist, as he would in the case of any other scientific anomaly—Michelson and Morley's first demonstration that the earth does not move through the ether, or Planck's first demonstration that energy is emitted in discrete quanta—that a single research by a single pair of investigators on a single pair of boys must be further confirmed by other investigators with other subjects, and some kind of working hypothesis must be elaborated which will not merely coordinate the new data with accepted

[1] The Welsh members of my own family and of my own college at Oxford (which was predominantly Welsh) have always insisted that such powers are normal rather than unnatural or paranormal; and if (they say) their own fellow-countrymen sometimes represent them as exceptional and mysterious, it is because that is the way the Sassenach likes it.

psychological theories (modified if necessary), but itself be capable of objective verification: only then can the results be accepted at their face value.

After all there are several manifest loopholes on which a stubborn sceptic could easily seize. Here are two Welsh youngsters, working at a monotonous task for a couple of grown-up visitors, coming from a distant city and speaking a foreign tongue. They are offered prize-money for high scores; they refuse to work in separate rooms with the door between them shut; and they are not averse to deliberate trickery. Does not such a situation frankly invite a theory of fraud ? Nevertheless, before the reader jumps to any such conclusion, he ought in all fairness to examine Dr. Soal's detailed report, and then explain how he himself imagines so prolonged a hoax could have been sustained in the face of the various counter-measures taken to defeat it. I doubt if he will succeed[1].

IV. The Planning of Future Research

The Basic Assumptions. To the majority of psychologists research on telepathy and kindred phenomena has appeared an unprofitable pursuit for two main reasons. First, there is at present no known method of eliciting the phenomena at will in order to analyse them experimentally. Secondly, there is no known hypothesis which will account for such phenomena and bring them into line with the postulates and principles of modern science[2]. Yet in the history of science it has frequently been just these two baffling features in out-of-the-way phenomena that have roused the efforts and the ingenuity of those persistent enthusiasts who in the end have revolutionized the views accepted by their contemporaries.

Dr. Soal himself believes that it would be premature to set about constructing any new theory. His own motto is, *Hypotheses non fingo.* "Instead of wasting time in sterile hypotheses", he says, "it will be far better to carry out more experiments to discover what pieces of the puzzle are missing." And he ends his last chapter by lamenting that "Glyn and Ieuan have taught us little or nothing about the nature of telepathy".

But here, I fancy, he is too modest. Scattered in the course of his comments are numerous incidental points which should enable us to take the first provisional steps. In any case, before we embark on further experiments, it is to my mind essential that we should get our provisional assumptions explicitly stated, and then formulate alternative hypotheses, which our investigation must be expressly

[1] The editors of the *Journal* would welcome criticisms and suggestions from readers who may have studied the work of Dr. Soal and Mr. Bowden. They understand that these writers themselves would gladly receive offers from anyone who would be prepared to demonstrate an alternative procedure that could produce similar results under conditions similarly controlled.

[2] Professor Gardner Murphy believes this second objection is unwarranted. "Unknown to most psychologists," he says, "several coherent theories have already been developed" (*Human Potentialities*, 1958, p. 295). With that I fully agree; but my point is rather that none of these theories (at any rate none with which I am acquainted, least of all those of Carington, Rhine, Thouless and Wiesner, and Gardner Murphy himself) are explicitly related to the current scientific background or furnish observable corollaries capable of being verified or falsified. They are *internally* coherent; but no more. That is why there are 'several', each different from the rest.

planned to test. Experiments which *seem* to be designed with no underly assumptions and with no guiding hypothesis nearly always owe their design to some tacit hypothesis at the back of the investigator's mind; and, as often as not, that hypothesis is too vague or too remote to be fruitful as a guide. It is largely this attempt to collect data in and for themselves, with no reference to broader psychological issues, that has caused the phenomena of 'extra-sensory perception' to stand apart in an anomalous world of their own. And certainly, since Dr. Soal's new work, added to the evidence already available[1], goes far to establish the *fact* of 'extra-sensory perception', it is high time to start systematic investigations on its *nature* and *modus operandi*.

The problem of providing some intelligible explanation of such phenomena can be attacked by two opposite and complementary lines of approach: we may designate them *causal* hypotheses and *epistemological* hypotheses respectively. Hitherto theories of the former type have been almost exclusively sought.

A. *Causal Hypotheses.* In the physical sciences the search for causes as such has now been almost entirely abandoned; but in psychology it still seems to be widely accepted as an indispensable element in any scientific methodology. Moreover, owing to the prevalence of a materialistic type of behaviourism, the causes sought for have been conceived for the most part in terms of the out-of-date mechanistic doctrines of nineteenth century physics. Thus, as we have seen, the two characteristics of 'paranormal' phenomena which make them so difficult for the scientist to accept are (*a*) that they imply relations between two distinct events which are in no way connected by any continuous causal chain, but seem separated, as it were, by a 'spatio-temporal gap', and (*b*) that in certain instances (e.g. pre-cognition) the later event appears to be the cause of the earlier.

There seem to be three principal ways in which investigators in this field have endeavoured to close the gap by inserting some hypothetical chain of cause and effect operating in the forward direction: (i) the physical, (ii) the biological, and (iii) the psychical. It is the last that falls more especially within the province of the psychologist. Dr. Soal, however, is inclined to reject it. To think along psychological lines, he believes, "is waste of time: it seems far more likely that we shall have to resort to biology and the physics of brain-functioning for any light on the mystery". Let us therefore consider these first.

(i) *Physical Theories.* With those who invoke "the physics of brain functioning" the most popular theory at the moment is a revival of one due

[1] In *Extra-Sensory Perception after Sixty Years* (1940) Dr. J. G. Pratt and his four co-authors have examined 142 published reports on ESP, and find six which in their view adequately excluded all alternative modes of explanation. To these should now be added the results obtained by Dr. Soal with Basil Shackleton and Gloria Stuart (S. G. Soal and F. Bateman, *Modern Experiments in Telepathy*, 1954). Together they yield strong evidence for telepathy and for pre-cognition (in the case of 'displacements') and for pure clairvoyance (though here the evidence is less convincing). The possibility of psychokinesis seems far more doubtful: Rhine's review of the subject (*The Reach of the Mind*, 1947) so often misquotes, or at least exaggerates, the results of other investigators (e.g. Smith, Gibson, Hilton and Baer) that one is tempted to wonder whether at times his other reports may not unconsciously suffer in much the same way.

originally to Sir William Crookes, namely, that telepathic communications are conveyed by "an outward radiation from the brain of extremely short and very penetrative electromagnetic waves"[1]. At first sight this seemed to receive some measure of support from the discovery of rhythmic changes in the electric potential at various points in the brain itself. However, when we consider the actual magnitudes involved, these more recent results tend rather to confute than to confirm any explanation along those lines. The alpha waves have a frequency of 8 to 13 cycles per second: they would therefore have a wave-length of about 30 000 kilometres. Their average amplitude is only 50 microvolts: a few centimetres from the surface of the head they would be incapable of affecting the most delicate apparatus that could be manufactured to detect them. It is therefore difficult to believe that they could influence another brain at a distance of four or five feet; and at 160 feet such an effect seems quite incredible[2].

Rhine himself considers that, as a method of connecting "new facts and old laws along lines of present physical theory" the "radiation hypothesis would be intellectually delightful"; but in the end he too decides that the "distance data" exclude it[3].

Nevertheless, he contends that, since "it is an axiom that physical processes cannot be regulated except by the expenditure of energy," we are "faced with the need for another order of energy, not radiant". Its precise nature, he thinks, is "a clear-cut problem for the physicist". In this vaguer form, however, the emanation theory ceases to be a testable hypothesis, and becomes little more than

[1] The conception has been widely popularized by Upton Sinclair's book entitled *Mental Radio* (1930, Preface by McDougall), and put forward afresh by several physicists (e.g. by J. D. Vernon, *The New Scientist*, V, 1959, p. 197).

[2] Direct attempts to pick up telepathic signals by means of appropriate electrical equipment have been made by von Ardenne (*Zeitschr. f. Technische Physik*, IX, 1928, pp. 288–9) who found evidence only of muscular, not of neural activity, and by F. Cazzamalli (*L'Energia Elettricia* XVIII, 1941, pp. 28–42), who believed that he obtained some slight indications of such signals: but I gather that few physicists would consider the methods used sufficiently reliable to demonstrate the inferences drawn.

[3] If we grant the rather fanciful suggestion that a particular brain might (as it is sometimes put) be 'tuned' to pick up such waves, then I doubt whether the 'distance argument' is as convincing as is ordinarily supposed: we could equally well imagine that such a brain could compensate for increased distance by means of some amplifying device. However, I know of no neurologist who would dream of entertaining this type of theory for one moment. Professor Eccles, an eminent neurologist and a Fellow of the Royal Society, has attempted to overcome the difficulty by supposing that what emanates from the transmitter's brain is not some 'electro-magnetic wave', but some 'psycho-kinetic influence': this is supposed to be capable of triggering off any neurones in the percipient's brain that are in a critical state of sensitivity (J. C. Eccles, *Nature*, CLXXV, 1951, pp. 516f., and *id*, *The Neuro-physiological Basis of Mind*, 1953). But, since we know less about psychokinesis than we do about telepathy, such a hypothesis only explains *ignotum per ignotius*. G. N. M. Tyrrell (himself an electrical engineer) points out a still more serious objection: transmission by electro-magnetic radiation would require the transmitter's thoughts to be first encoded, and then decoded by the agent. But this in itself seems by no means fatal: for, as ordinarily interpreted, the normal processes of perception and of volitional movement also involve a decoding and encoding of the signals conveyed to and from the brain through the usual neural channels.

the gratuitous use of a technical term from modern physics in a context where there is no evidence that it is in any way relevant[1].

(ii) *Biological Theories.* Rhine combines his theory of a new form of energy with the argument, put forward independently by several writers, that "ESP is a biological phenomenon, having survival value for any species possessing it". In keeping with this theory we might suppose, as indeed has often been maintained, that primitive races inherit these extra-sensory powers as part of their natural endowment, but that with the development of surer methods of securing information it has atrophied through disuse[2]. This would explain why children from remote rural districts, like Glyn and Ieuan, still apparently possess it, while more sophisticated town-dwellers such as university students seem almost wholly devoid of it.

Rhine goes on to suggest that such a hypothesis might well account for much that is puzzling in the homing, migrating, food-gathering, mate-finding, and migratory activities of certain insects, birds, and mammals. "If it has not been more clearly observed, that," he believes, "is simply because it has not been looked for"; and more recently Dr. J. G. Pratt and Professor A. C. Hardy have advanced somewhat similar arguments, though in far more cautious terms. In point of fact, however, several naturalists have investigated the problem with this possibility expressly in mind; and most of the evidence points to much simpler types of explanation[3]. Moreover, even were socalled paranormal phenomena actually part of the normal biological heritage of many surviving species, that of itself would not explain their mode of operation[4].

[1] *Extra-sensory Perception* (1934), pp. 118f. Fifty years ago several psychologists attempted to interpret consciousness itself as a form of energy. The first to put forward such a suggestion was Professor Lloyd Morgan. McDougall at one time postulated a special sort of ' psychophysical energy '. But the clearest exposition of this view is that of W. P. Montague who thinks it provides a feasible solution for what he calls ' the most perplexing of philosophical problems ' (' Consciousness a Form of Energy ', *Essays in Honor of William James*, 1908, pp. 103–134). I believe each of these writers eventually abandoned the conception as quite untenable in the light of subsequent criticism.

[2] Cf. J. C. Dorsainvil, *Voudou et Nevrose* (1931) and B. J. F. Laubscher, *Sex, Custom, Psychopathology: A Study of South African Pagan Natives* (1937): also M. Pobers, ' Psychical Phenomena among Primitive Peoples ' and J. Langdon-Davies, ' Extrasensory Perception among Peasant European Populations ' *ap.* G. E. W. Wolstenholme and E. C. P. Millar (ed.) *CIBA Foundation Symposium on Extrasensory Perception* (1956), pp. 102–130. It would, I think, be generally agreed by most investigators that the evidence so far obtained is by no means convincing: yet the observations recorded, particularly when taken in conjunction with Dr. Soal's own findings, strongly suggest that inquiries in undeveloped countries should be systematically carried out before the advance of civilization impairs the use of whatever capacities such peoples many perhaps possess.

[3] See G. V. T. Matthews, *Bird Navigation* (1953) and G. Kramer, *J. Ornith.*, XCIV, 1953, p. 201f, XCVI, 1955, p. 173f. The remarkable studies of Professor von Frisch on honey-bees are sufficient to warn us of the unsuspected possibilities of previously ' unknown senses ' (e.g. sense organs in the antennae to air-borne molecules of water) and of unknown modes of using ' known senses ' (e.g. in the interpretation of the bees ' dance as a means of communicating the direction and distance of food).

[4] An attractive theory from a biological standpoint would be a modification of the ' instrumental ' or ' transmissive ' hypothesis put forward by James. Instead of supposing that the brain generates consciousness as a lantern generates light, we might suppose that it focuses consciousness in the direction requiring attention much as the lens of a telescope or microscope focuses light where necessary and excludes unnecessary light. Professor Broad seems to incline to a view of this sort:

(iii) *Psychic Theories.* A few investigators have suggested the possibility of some specifically psychic form of causation. Bertrand Russell, it may be remembered, sought to account for our memories of the past by postulating a direct causative effect exerted by earlier mental experiences on later mental experiences, without assuming any continuously persisting physical 'traces' left in the structure of the brain. Similarly Professor Broad has suggested "an immanently and purely psychical causation" acting backwards in time, whereby a later mental experience might exert some kind of causative influence on earlier mental experiences, and so give rise to pre-cognition[1]. It would be extremely difficult, but not, I fancy, wholly impossible, to devise experiments which could put such theories to the test. Meanwhile, however, they remain entirely *ad hoc* assumptions, and, as such, highly suspect.

The more popular types of psychological explanation are nearly all based on the old-fashioned dualistic doctrine of an interacting body and mind. The mind is conceived, in scholastic terms, as an individual substance like the body, differing from the body by being of an immaterial rather than material nature, and endowed with unique causal properties, namely, its various 'faculties'. In order to account for the paranormal phenomena, it is then assumed that the mind is divided into at least two portions or layers—a conscious mind and what Myers calls the 'subliminal self' (i.e. a self existing or functioning below the threshold of the conscious mind). Recently endeavours have been made to relate subliminal self with the unconscious mind as described by Freud and Jung. The phenomena of telepathy, clairvoyance, pre-cognition, mediumship, and the like are usually grouped together as manifestations of a single 'faculty' possessed solely by the 'subliminal self'—a capacity (as Tyrrell suggests) which is "something very remote from the faculties of ordinary experience"[2]. This peculiar power is then conceived as operating, not only from mind to body or from body to mind, but directly from one mind to another.

Such a hypothesis is open to all the familiar criticisms urged against the doctrine of faculties and the dualistic theories of body and mind. Moreover, if the faculty postulated (the 'sixth sense', as some writers have called it) were just a simple but hitherto unrecognized cognitive ability, and especially if it were (as Rhine suggests) a 'biological capacity' presumably dependent on additive genes, then we should expect the individual variations in its amount to exhibit

" the suggestion is that the function of the nervous system and sense-organs is *eliminative* not productive: each person is potentially capable of perceiving everything that is happening everywhere in the universe; the function of the brain is to protect us from being overwhelmed and confused" (*op. cit.*, p. 23: I would add *concentrative* to *eliminative* in describing the brain's function).

[1] 'Philosophical Implications of Foreknowledge', *Proc. Arist. Soc.*, p. 116.

[2] G. M. N. Tyrrell, *Science and Psychical Phenomena*, 1938, p. 116. I believe F. W. H. Myers was the first English writer to mention Freud and his theory of the unconscious. Freud states emphatically that " it appears no longer possible to disregard the investigations of so-called occult phenomena" (*Gesammelte Werke*, XVII, 1921, p. 25): in his view ideas which have a strong emotional tone might, as they emerge from the repressed or unconscious state into consciousness, be transferred from one person to another particularly when the latter was asleep: (cf. *Collected Papers*, V, 1925, p. 158).

a frequency-distribution similar to that described by the normal curve, i.e. a distribution of Type IV (a bell-shaped curve, moderately asymmetrical, approximating to the normal). As it is, the surveys carried out by Dr. Soal at University College and through the B.B.C. suggest (if I am not mistaken) that the distribution is rather of Type VI (nearly if not quite J-shaped) or possibly even Type XI: by far the commonest individuals are those who have little or no ability, and those showing discernible amounts vary considerably, the extreme case being decidedly rare. Such a distribution arises when the factors involved are not additive (as in curves for abilities) but multiplicative (as in curves for output or income)[1], for example, whenever, in addition to the basic capacities, complicating determinants like motivation and various environmental tendencies (e.g. social requirements)are involved. This of itself suggests important lines of investigation.

So far as any basic 'faculty' is concerned the first thing for the theorist to prove would be the presence of a definite 'group factor' underlying all the different paranormal manifestations; and, instead of splitting the mind into two (or, as some would have it, three) distinct compartments, most factorists would prefer the hypothesis that the whole psychophysical personality is organized as a hierarchical structure. Such a conception could be far more readily adapted to explain the phenomena in question[2]. But to my mind the evidence for this broader theory should be sought rather in the study of normal mental processes than of paranormal.

However, in the present state of knowledge, it seems quite as futile to seek causal explanations for paranormal processes as it would be for any other types of conscious activity. Accordingly, the first requisite is not to think up conceivable modes of causation, but to determine far more precisely what is the specific nature of the processes we want to explain. For the moment we must postpone all attempts at constructing some complete and well-rounded theory, which will cover every aspect at once, and proceed instead step by step, taking two or more mutually exclusive but extremely simple hypotheses at each successive stage. Hitherto, as we have seen, most investigators have treated paranormal manifestations as though they formed a self-contained collection of phenomena to be examined and explained in isolation, as something quite unique. This is clearly a mistake. Whatever else they may be, paranormal manifestations are all modes of consciousness, or at least of conscious behaviour. It follows that we must begin by formulating hypotheses which may have a meaning in the wider context of conscious processes in general.

[1] See Burt, 'The Distribution of Intelligence' *Brit. J. Psychol.*, XLVIII, p. 161f., and *id.*, 'Ability and Income', *Brit. J. Educ. Psychol.*, XIII, 1943, pp. 94f. To psychologists, the most familiar instances of such curves are those for income (fitted by 'Pareto's equation') and Allport's curves for 'conforming behaviour'.

[2] If we like to assume that the universe also has a hierarchical structure, it would then be tempting to suppose that, contrary to the assumption so staunchly maintained by the traditional individualism of the West, there are degrees of independence and of interdependence between man and his environment, as Gardner Murphy has so often and so eloquently urged (see, for example, his *Human Potentialities*, 1958).

B. *Epistemological Hypotheses.* Once we have grasped the fact that many of the features which seem so puzzling in paranormal phenomena characterize nearly every form of conscious process, they begin to lose much of their incongruous and sensational character: we are no longer tempted to throw them out as misfits in our general psychological scheme. After all, we readily admit that it is quite impossible to discover any 'continuous causal chain' linking the changes in the cortex of the brain with the everyday types of conscious experience associated with such changes. Here too there is a 'spatio-temporal gap'; and the familiar facts of normal sensory perception are as deep a mystery as the facts of so-called extra-sensory perception. Again, in purposive behaviour, an event in the future seems to determine events that precede it. If, with the behaviourists, we resolve to exclude any reference to sensory experience or to purpose, then we must expect to be baffled when we try to understand telepathy and precognition; but if we really mean to learn more about the distinctive nature of paranormal modes of consciousness, it seems plain we must relinquish a rigidly behaviouristic attitude, and fall back on the only way of studying conscious phenomena at first hand, namely, the old-fashioned method of introspection.

(i) *Cognitive versus Spatio-Temporal Relations.* A moment ago, while sitting at my desk, I saw above the houses opposite a flash of lightning succeeded by a roll of thunder. When by an effort of introspection I try to analyse these experiences in verbal terms, I find it altogether inappropriate to describe my awareness of such things as either a 'substance' or an 'attribute': to me at any rate it seems plainly to be a 'relation'—a quite specific if transitory relation between me on the one hand and the flash or the sound on the other. Logically it belongs to the same category as the relation of succession between the flash itself and the sound that followed. Just as I can sit *on* my chair, so I can be conscious *of* my chair[1]: in both cases what the two prepositions designate are relations. Moreover, I should argue that in all these experiences I am actually conscious of the relations. Just as I see that the lightning is *above* the roofs and not below, so I am aware of the fact that I *am* aware of the light and the noise. I cognize my cognition.

Now, unlike earlier writers who relied heavily on introspection—Spencer, James, Stout, Spearman and their various disciples—most contemporary psychologists scarcely mention relations and make little or no use of them in their expositions and arguments. Their concepts and their categories are for the most part those of the traditional methodology which endeavoured to explain all phenomena in terms of *substances* like matter or energy, which are located in Euclidean *space* and *time*, and are endowed with certain attributes or *properties*, such as the capacity to *cause* certain effects. In the days of scholastic science this simple conceptual scheme had the advantage of allowing the scientist to cast his arguments into the subject-and-predicate form, which until recently

[1] I do not mean to imply that the pronoun 'I' necessarily designates the same grammatical subject in both cases.

was the only type of rational argument that orthodox logic could handle. But, as the mathematicians and the physicists have abundantly shown, a mere syllogistic mode of reasoning is wholly inadequate. In modern science the logic adopted is essentially a calculus of relations[1].

Once we recognize that all cognition, whether normal or paranormal, implies a unique kind of relation, then the anomalies of which Rhine and others make so much begin to disappear. Rhine, as we have seen, conceives mind or consciousness to be a quasi-physical entity or substance—a kind of energy; and he therefore tends to express his problems in spatial terms. He asks what is the "reach of the mind in space", where are the "frontiers of the mind", how great is "the measurable force of mind", and how these paranormal experiences enable man's consciousness to "jump across the barriers of time". It is this way of envisaging the problem that leads to the puzzles and paradoxes about the way "ESP transcends the limitations of time and space". After all, every relation that is not itself a spatial or temporal relation, or (like cause and effect) defined in terms of such relations, "transcends the limitations of time and space". And this is as true of sensory perception as it is of extra-sensory. At midnight, towards the end of October, if I look due south I can see something that was happening over 2 million years ago at a distance of 13 million million million miles—the spiral galaxy in Andromeda as it was at that date. In memory I can cognize many events, like Queen Victoria's jubilee, which happened last century; and there is no *logical* impossibility in the suggestion that I might possibly cognize events that will not happen until next century. If I cannot perceive them, I can at least think about them; and in some cases I can think about them with some prospect of being correct: a few simple calculations, for example, tell me that the next transit of Venus will occur in June, A.D. 2004: one of my women-students, who is also an amateur astronomer and enjoys eidetic imagery, tells me that she "visualizes the transit with all the vividness of a precognition". The only difficulty, therefore, about the recorded instances of clairvoyant predictions is that we cannot at present suggest any satisfactory grounds for their apparent accuracy, as we can in the case of the astronomical forecast.

(ii) *Thought Processes versus Perceptual.* We may then safely assume that consciousness (in the sense of being conscious *of*) is neither a substance like energy), nor yet a property (like being alive), but a relation, and that it is one of the many types of relation that transcend space and time. My next point is that, as

[1] See E. Cassirer, *Substanzbegriff und Funktionsbegriff*. It will be noted that the view of consciousness put forward here seems to rule out Dr. C. C. L. Gregory's interesting suggestion. He holds that what we ordinarily call 'objects'—i.e. substantial entities like chairs, molecules, stars, or at least the particles of which they are composed (electrons, protons, etc.)—are 'images of great mass and energy': similarly all mental images are really substances of minimal mass and energy, having location and duration like ordinary objects ('A Cosmological Approach to a Theory of Mental Images', *Proc. Soc. Psychic. Res.*, LII, 1958, pp. 41–4). But a mental image is a 'content' which we directly cognize. If electrons and protons are images in this sense, where is the consciousness that cognizes them? On my theory, where there is no cognitive relation, there is no cognitive image.

introspection shows, there are at least two distinct types of cognitive relation. In nearly every language except our own these have different names—*nosse* and *scire*, *kennen* and *wissen*, *connaître* and *savoir*. Thus, when you or I perceive a card that Dr. Soal holds up we not only cognize the white sensory patch that we *ken*; we also cognize a material thing that we *wot*. The white sensory patches are peculiar to each of us; but the actual thing we perceive or think of is the same for you, for me, and for Dr. Soal. The transitory sensory elements of which we are immediately aware I shall call 'contents'; the more permanent and public elements, like the card, which we cognize only indirectly, I shall call 'objects'[1].

This distinction suggests the first pair of alternatives for the investigator of E.S.P. to consider and test: assuming that paranormal experiences are essentially manifestations of some kind of cognitive relation, is that relation one of (*a*) *direct* cognition or (*b*) *indirect*? In other words, are such experiences anomalous modes of *perception* in which the bodily sense-organs are not concerned, or are they modes of *thought*? The fashionable phrase 'extra-sensory perception', and such terms as 'clairvoyance' (or 'second sight'—seeing clearly what normally would be hidden) and 'telepathy' (receiving experiences from a distance), all presuppose, without a shred of proof, that the former alternative is the correct one[2]. They tacitly imply that in telepathy the 'percipient' (Glyn, for example) perceives the same sensory content as that perceived by the 'agent' (Ieuan), and that in clairvoyance (e.g. naming the cards in order when the pack has been is shuffled and then left untouched) or in precognition (e.g. dreaming, as a Dutchman did, that the number 3,684 would win the 1937 State Lottery), the percipient perceives or imagines a sensory content of visible shapes and colours, with no stimulation of the retina such as would occur in ordinary sense perception. I suggest that in all such cases it is far more probable that what we have to deal with is, not 'percept-transference' or even 'thought-transference', but rather a mode of implicitly guiding or directing thought so that the so-called 'percipient' *thinks* of the right object or the right situation at the right moment.

As far as telepathy is concerned, this interpretation seems in some measure supported, not only by the introspective comments I have myself secured from the few 'percipients' that have come under my notice, but also by several observations incidentally recorded by Dr. Soal and Mr. Bowden in their work with the two

[1] The distinction between these two modes of knowing—'knowledge by direct acquaintance' and 'indirect knowledge about'—are familiar to psychologists from the account given by James (*Principles of Psychology*, I, p. 221f.) and to philosophers from Bertrand Russell's various writings (e.g. *Problems of Philosophy*, chap. V). The distinction is an ancient one. It was clearly stated by St. Augustine (*De Magistro*); and is perhaps most instructively discussed by Twardowski (*Zur Lehre vom Inhalt und Gegenstand der Vorstellungen*, 1894). By far the best known formulation is that of Meinong (*Uber Gegenstände Höherer Ordnung*, 1899).

[2] The standard definition given in the *Journal of Parapsychology*—a definition explicitly accepted as 'excellent' by the CIBA symposium—namely "response to an external event not presented to any known sense"—appears to me most inappropriate. Much of the anticipatory behaviour of an intelligent person in everyday life consists of "responses to external events not presented to any known sense"; and so long as the sensory functions of the semi-circular canals remained unknown, the responses to changes in the position of the head would have been classed as modes of E.S.P., had this definition been adopted.

Welsh boys. For instance, like most subjects who make successful scores in experiments of this kind, Glyn is apparently unconscious of any sensory stimulus or similar external influence; nor is there any hint of hallucinatory percepts; and, unlike most others, Glyn apparently possesses little or no visual imagery, whether eidetic or otherwise[1]. However, more direct and systematic confirmation is needed before such a hypothesis can be regarded as established.

As to precognition and clairvoyance, the hypothesis as worded would cover these also, were they clearly proved. The chief problem here would be to state what in such cases *is* the 'right' object or situation; otherwise we may be arguing *ex post facto*. The evidence available suggests that, even in 'pure' clairvoyance (if it exists), the 'right' object or situation is selected in advance (unconsciously if not consciously, and in general terms if not specifically) by some *person*: this may be either the percipient himself or some other individual, e.g. the agent[2]. However, this again calls for further confirmation.

(iii) *Synthetic versus Analytic Processes.* If, then, the process involved turns out to be some kind of thought process, the next step will be to determine whether it is an 'analytic' or a 'synthetic' type of process. The evidence available suggests that in all probability it is a process of the latter type—a kind of 'intuition' or 'insight' (in the Gestaltist's sense), but this too is a problem on which far more psychological research is needed.

Unlike perceptual processes, thought processes, particularly those of the 'synthetic' type, are apt to be highly irregular in their nature and occurrence. They are more spontaneous, more creative, and more difficult to predict than

[1] I suggest that the common notion that what is transferred is a *visual* percept or 'image' is due to two main reasons. First, the material generally used (drawings, concrete objects, diagrams on cards, etc.) is of the type which calls for designation in visual terms. Far more experiments with non-visual concepts are needed: for instance, our 'agent' might be (say) a German who knows no English, and we might give him ideas to transmit that evoke no obvious visual images and are expressed in his own native tongue—e.g. *Pflicht, Tugend, Genauigkeit*, etc., the 'percipient' would be an Englishman who is ignorant of German, and he would be given cards expressing the same ideas in English—*Duty, Virtue, Accuracy*, etc.; our aim would then be to see how many correct matches the 'percipient' makes. But secondly, the most successful 'percipients', and the processes on which they rely, appear to be nearly always of a synthetic, non-intellectual type; and in such cases the 'thoughts' are apt to be re-cast into visual form. The most familiar examples are those of daydreams and nocturnal dreams. Thus, even when the ideas to be transmitted have a purely verbal form in the mind of the agent, they seem usually to be experienced as visual symbols by the 'percipient'.

[2] Here the most striking examples are anecdotal and therefore chiefly visual: several are well attested: e.g. Swedenborg's vision of the fire at Södermalm and his announcement that a lost receipt was to be found in a certain secret drawer in the Dutch ambassador's house (both discussed by Kant), Rider Haggard's vision of the death of his favourite retriever, or that of Dr. Osty's hypnotized subject who described in great detail a corpse lying in a thicket where it was subsequently found. But it is impossible to evaluate the odds that these are not after all just unusual coincidences; nor is it easy to rule out telepathy. The best statistical evidence is that obtained by Pratt and Pearce (*J. Parapsychol.*, I, 1937, pp. 70–80.) Among the six 'satisfactory researches' cited in *Extrasensory Perception After Sixty Years* (see p. 74 above) this series stands out as one of the two most carefully conducted: it was carried out with the usual type of Zener cards. Pratt selected each card, but did not look at the face. Almost equally convincing are the experiments carried out by D. R. Martin and F. P. Stubic (*J. Parapsych.*, II, 1938, pp. 23–30, IV, 1940, pp. 159–248). To all these the hypothesis suggested above would apply quite neatly, as well as to the numerous laboratory experiments of the 'down-through' or of the 'displacement' type.

processes of perception. These characteristics, found in all the commoner forms of thinking, would go far to account for the fact that paranormal processes are so frequently described as "spontaneous, creative, and difficult to predict": in these respects therefore they are in no way 'paranormal'[1]. Indeed, it is just because their occurrence is so irregular, and depends on so many different variables, that the continual demand for single, definite and repeatable types of observation, produced at will under appropriate experimental conditions, has been so hard to satisfy; and a statistical analysis of large groups of data seems to offer the only satisfactory method of demonstration.

Persons versus Situations. But, it may be said, there is a world of difference between merely *thinking* that this card (or the next) will be an Elephant, and being *right* when one thinks it will be an Elephant: does not this imply some unique sort of power or capacity on the part of the 'percipient'? Here, I think, we approach the crucial problem. Yet the question as ordinarily put is wrongly phrased. It implies the outworn scholastic logic which assumes (as we have already seen) that solutions are to be expressed in terms of individuals and their capacities, properties or powers—a logic long ago caricatured by Molière, but still persistent in semi-popular science. "It's a gift," says the young nurse of the surgeon's diagnosis in *Tell Me Doctor*, "the true clinician's gift! He just looks at each patient in turn with his X-ray eyes, sees all that's hidden inside her, and then says—'That is an appendix; that's a gastric ulcer; and that's a carcinoma—inoperable I fear.' And he's always right; what's more he knows he's right!" In the same way it is widely supposed in paranormal research that the successful clairvoyant or telepathist is a peculiar kind of person with a peculiar kind of 'gift'—some one standing apart from the rest of the community, as the 'genius' is supposed to do, in virtue of some exceptional faculty or flair. But in nearly every well-attested study even the best of percipients is frequently wrong, and certainly does not know when or whether he 'is right': he cannot even say which of his guesses is most likely to be right. All that can be safely asserted is that his special pattern of guessing includes a higher proportion of correct guesses than is usually the case. And, as the details recorded by Dr. Soal and Mr. Bateman indicate, such patterns seem to be associated primarily with particular types of psychological *situation*, and not just with particular types of *person*.

No doubt the characteristics of those who participate play an important part[2]. But much more than a suitable choice of person is required. Dr. Soal's

[1] Tyrrell, loc. cit. *sup.* It is the fashion to stress the resemblances between paranormal phenomena and certain *exceptional* types of thought process—the inspirations of the artistic genius, the ecstatic states of the mystic, the predictive insight of the religious prophet; and not uncommonly these latter are also attributed to the workings of the 'unconscious mind' or the 'subliminal self'. In point of fact, it is not the processes themselves, but the results of the processes, that are exceptional.

[2] As I have already suggested, percipients whose mode of thinking is synthetic rather than analytic appear to be the most successful; but certain temperamental characteristics seem essential as well, and these imply a special attitude, or at least a special responsiveness, to personal situations. In terms of my own classification of emotional types (*Character and Personality*, VII, pp. 238f., 285f.) the 'stable extravert' is usually more successful than the 'unstable' or the 'introvert'.

experiments, as we have already seen, bring out several other conditions needed for success. Perhaps I may venture to indicate, mainly from my own limited experience, a few further points that appear to me worth following up.

To begin with, it seems essential to make the whole situation as interesting as possible. Dr. Soal's substitution of coloured representations of animals for the abstract geometrical patterns used by Rhine and his followers is undoubtedly a step in the right direction, particularly where the participants are children. But the interests aroused need not be obvious or conscious; and I would suggest further work with a much wider variety of pictures and symbols, particularly those likely to arouse emotion and attention, or to touch on repressed complexes (sex, love, father or mother, birth, death, and the like). Mood we have seen to be supremely important; and most psychologists engaged in individual testing learn, so to speak, a clinical or 'bedside' manner, calculated to set up and sustain an appropriate attitude in their testees. Here again Dr. Soal has singled out one valuable device by demonstrating the remarkable effect of a financial incentive.

Another method of influencing mood and emotional background which deserves further trial is through the use of drugs. Stimulants (e.g. caffeine) are said to raise the scores, "presumably," it is suggested, "by facilitating concentration"; sedatives (e.g. sodium amytal) are said to lower them by blunting effort and emotion[1]. Little has been done with drugs having more specific effects on the nervous system (e.g. nitrous oxide, mescalin, lysergic acid, and the innumerable synthetic novelties now available). The ideal situation seems to resemble that which the psychoanalyst seeks to maintain in dealing with his patients—a condition which combines inattentiveness to thoughts and emotions that are irrelevant with a quiet concentration on those that are essential, and a confident desire to achieve success—a condition which (in my meagre

This seems borne out by the recent investigations of G. F. Schmeidler and R. A. McConnell (*E.S.P. Personality Patterns*, 1958). They found that "the dilated ambiequal (well-balanced) subjects had scores above the coartative (overly repressed)", and that "a passive unreasoning acceptance of impressions is conducive to superior performance". The "aggressive" usually failed. In particular those who accepted E.S.P. as a genuine possibility had higher scores than those who rejected it.

Moreover, as the results obtained by Dr. Soal with Glyn clearly show, the personality of the agent may be almost equally important. This is also indicated by the result of an ingenious series of experiments reported in his earlier book. In these, *two* agents simultaneously transmitted *different* messages: in the course of six runs carried out under these conditions Shackleton's scores rose well above expectation in every case with one agent, and tended to fall rather below expectation with the other. Curiously enough, the type of personality requisite for a successful *agent* has never been systematically examined. Rhine, however, suggests that an agent or experimenter who adopts a highly scientific and critical attitude is liable to fail, and even suggests that with students Soal's manifest demand for rigorous conditions may have exercised an "inhibiting effect."

[1] J. B. Rhine, *Extrasensory Perception*, p. 127f. Mr. T. L. Jackson, however, tells me that in his experience 'tranquillizers' in mild doses tend to raise scores, "presumably by quietening emotional stress and irrelevant fancies". It would be interesting to learn whether psychoanalytic sessions, either with patients or with trainees, would not prove an effective way of discovering persons with telepathic abilities: for suggestive instances, see W. Gillespie, *ap. Psychoanalysis and the Occult*, p. 373f; also H. Deutsch, *Imago*, XII, p. 4188 and S. Hollis, *Imago*, XIX, p. 529f.

attempts) has been best obtained under mild hypnosis[1]. However, at this stage of the work what seems most urgently needed are not simply surveys to discover likely persons, but surveys to discover likely types of situation.

But I need not prolong my enumeration of the many problems, hypotheses, and alternative procedures which Dr. Soal's results would seem to suggest for future research. There are, however, one or two methodological cautions of a more general nature that need to be indicated by way of conclusion.

A Field Theory. Although in the initial stages of an inquiry, it is desirable to start with hypotheses so framed as to be mutually exclusive, yet it often turns out that both are in some measure true: each may hold under certain conditions; each may express partial aspects of the truth which are complementary to each other. In the sphere of psychical research two all-embracing theories are constantly cropping up which are generally treated as mutually exclusive—the pluralist and the monist. Pluralists, like McDougall and Rhine, assume that individual minds (some of which may be disembodied) are self-subsisting substances, existing in a universe which is otherwise empty of mind; these 'psychic monads' can enter into causal action with certain material organisms, usually their own bodies or brains (a type of action which is thus essentially 'psychokinetic'); but they are also capable both of a similar psychophysical interaction with other brains (and possibly with inorganic bodies) and of a direct telepathic interaction with other minds. Monists, on the other hand, like F. W. H. Myers, prefer the Spinozistic conception of one all-pervading mind, " hidden, like the deep bed of the ocean, but thrusting peaks into the upper air, which seem to be islands, but are not, since below the surface they are united "[2]. Their challenge to the pluralists is that of Tennyson's *Ancient Sage*:

> Thou canst not prove that I who speak with thee
> Am not thyself in converse with thyself.

Sometimes the analogy is changed; and 'the one subliminal mind' is itself compared to an 'all-enveloping sea, in which each individual moves and has his being'. Thus, for McDougall the mental universe consists of many conscious atomistic *particles*; for Myers of a single psychic *medium* like the ether.

[1] A class-mistress of my acquaintance who carried out experiments with her pupils tells me she obtained highly significant results with two of her girls in the classroom, but failed when she sought to continue them after school hours in her own room at home where they evidently felt embarrassed; on resuming them under classroom conditions, she was once again successful. Schmiedler and McConnell believe that anything that tends to induce a " preoccupation either with *self* or with *process* forms a barrier to success; on the other hand, situations of emotional stress may at times be productive of psi effects ".

[2] Myers and many who worked with him were inclined to regard clairvoyance, and the higher manifestations of the 'subliminal self', as akin to the experiences of Eastern and Western mystics. At that date parapsychic speculations had acquired some measure of academic respectability, partly owing to the neo-Hegelian doctrine (upheld by T. H. Green and others) of the " one Eternal Oversoul spreading through many individual organisms ", and partly owing to the increasing interest in Neo-Platonic conceptions resulting largely from the studies of Plotinus and his followers by W. R. Inge and others at Oxford. Myers' pan-psychic hypotheses emerged more clearly in his talk than in his writings.

The difference is reminiscent of the old controversy as to whether the material universe consisted of numerous corpuscles moving in a *vacuum* or a single *plenum* varying solely in density—a controversy which started with Democritus and Aristotle, was revived by Descartes and Newton, and has endured to the present day. It is tempting, as I have elsewhere urged on rather different grounds, to argue that in psychology as in physics both views might to some extent be reconciled by introducing the concept of a field. Certainly, if in the physical sphere we adopt a field theory, in place of the classical mechanics of 'particles in empty space' assumed by Rhine and Hansel, we should surmount many of the difficulties on which they lay so much stress—e.g. the problem of apparent action at a distance. This does not imply that we "are trying to over-ride the facts of mechanics and the laws of physics", as Hansel and others appear to suppose: on the contrary, by invoking a modified field-theory, of the type developed in relativity dynamics by Lorentz and Einstein[1], and adapting it to the problems of quantum mechanics, contemporary physicists have been able to deduce all the familiar 'laws' of Newtonian physics, and to explain many of the anomalies as well, without importing the dubious concept of causation[2]. Similarly I suggest that in the psychical sphere we should substitute a field theory for causal theories, thus escaping many of the dilemmas which causal theories inevitably raise, and at the same time bringing both normal and paranormal psychology more into line with the cosmological theories of modern physics.

The term 'field' I use, not in the rather loose way in which it is sometimes used by 'topological psychologists', but in the stricter sense in which it is used by contemporary physicists. As conceived by Faraday and Maxwell, the essence of a field-theory consisted in supposing that the electric charge of a body lies not within the body, but in a spatial continuum around it. In much the same way we might suppose that the consciousness of a given individual, so far as it has location, resides in the environment around him rather than inside his head[3].

The introduction of yet another kind of field would in no way "conflict with the existing theories of science", since the physicist is always ready to

[1] According to Einstein's theory of 'special relativity', there can be no such thing as absolute simultaneity; and, it might perhaps be argued, this of itself would partly overcome the paradox presented by pre-cognition, where the event cognized appears by ordinary reckoning as something subsequent to, not simultaneous with, the act of cognition. As Eddington observes, once space and time are merged in a single continuum, "there is nothing to prevent a person being aware of an event before it happens; and an event might even 'cause' other events preceding it" (by the ordinary reckoning). Moreover, in one of his later attempts at a unified field Einstein proposes to add a new dimension to the more familiar concept of a four-dimensional space-time: this might permit us to adopt (were there adequate grounds for it) Broad's hypothesis of "a *second* time-dimension" ('Philosophical Implications of Foreknowledge', *Proc. Arist. Soc.*, XVI, Supp., 1937: see also F. Kaluza, 'Zum Unitätsproblem der Physik', *Preuss. Akad. Wiss.*, LIV, 1921, p. 966f., and A. Eddington's posthumous *Fundamental Theory*, 1946). I myself would suggest that we interpret these 'dimensions' as 'degrees of freedom', and so avoid the customary visuo-spatial conception that makes telepathy and allied phenomena so difficult to fit in.

[2] See C. Burt, this Journal, XI, p. 75f. and refs.; also G. Wentzel, *Quantum Theory of Fields* (1949) and E. Cassirer, *loc. cit. sup.*

[3] Kant indeed held that it was reasonable to say that a person's soul is present equally at every place where it would be natural to locate his sensory percepts.

postulate a novel type of field, provided it yields verifiable consequences; and the fields of the psychologists would have much the same formal structure as those of the physicist[1].

It is true that the psychic field would appear at first sight to be something quite disparate from the fields already recognized by physicists; but a similar disparity obtains within physics itself in the case of gravitational fields and electro-magnetic fields, which have so far defied all efforts to reduce them to a single 'unified field' or to a single set of coefficients describing the curvature of space-time.

Any two fields may interact. Whether they do so or not will depend on the particular conditions determining the so-called 'transition probabilities'. Any two fields may be wholly or partly isomorphic: thus the mental field belonging to one person might be in part a replica of that belonging to another person, and either or both might be in partial correspondence with a physical field. The so-called material structures we can regard as specific and quasi-local conditions limiting, rather than causing or producing, the active manifestations of the fields. Thus there may be fields within fields, the whole forming a hierarchical system— a 'world model' which would not only embrace the various aspects of the physical universe, but also coordinate the physical or non-conscious aspects and the psychical or conscious aspects in a single psychophysical cosmology[2]. In this way, as it seems to me, the pluralistic and the monistic hypotheses might be reconciled, and the apparent conflict between physical science and the results of psychical and psychological research be overcome.

Nevertheless, until we have more precise knowledge as to the nature of the paranormal processes, it would, in my view, be premature to attempt a formulation of so comprehensive a hypothesis in any detail. My purpose is merely to warn research workers and their critics against begging the whole question from the outset by thinking (as most of them apparently still do) in terms of a universe consisting primarily of tangible objects, visualized as occupying three-dimensional space, and obeying the simple principles of Newtonian dynamics. "As a result

[1] In Maxwell's theory the electro-magnetic field was pictured as a concrete elastic medium— the ether—in a state of internal tension. Such a state and its consequences can be described mathematically by using certain matrices known as 'strain tensor' and 'stress tensor'. The ether of the Victorian physicist, however, has proved as elusive as Mrs. Harris; and Einstein finally convinced us all that 'there's no sich crittur'. In his field theories a 'tensor' no longer implies an actual medium; it is merely a particular type of algebraic matrix which is (i) a function of a matrix of coordinates and (ii) obeys the same rules of transformation as convariant and contravarian vectors. Now factor-analysis, which the psychologist uses to describe the 'vectors of the mind' is also an application of matrix theory: indeed some factorists prefer to express their arguments by the aid of tensor-notation. And, when applied to the study of 'psycho-dynamics', the whole of factor analysis can be interpreted as a kind of field-theory.

[2] The most interesting attempt at psychophysical cosmology of the hierarchical type is A. N. Whitehead's 'philosophy of organism' (cf. *Process and Reality*, 1930, and *Nature and Life*, 1934). Among contemporary philosophers metaphysics and *Naturphilosophie* are out of fashion; but a still more recent speculative theory has been outlined by an astronomer who has explicitly sought to incorporate the findings of psychical research (C. C. L. Gregory and A. Kohsen, *The O-Structure*, 1959).

of the revolutionary changes both in physics and in epistemology,'' says Heisenberg, '' our conception of matter has become far less material, and our conception of mind far less mental: above all, we have learnt the perils of thinking in terms of visual diagrams and mechanical models, and the old scholastic logic ''.

Need for a Standardized Technique. Finally, in all the researches I have outlined, if only to forestall the kinds of objection so frequently encountered, it will be imperative that the most stringent conditions should be imposed on the conduct of the actual inquiry. Indeed, I would contend that, as in testing other types of mental process, it is urgently desirable to formulate an agreed and standardized technique. Dr. Soal's investigations have done much to indicate both what is feasible and what is necessary. May I, quite tentatively, put forward the following suggestions ?

(*a*) *Experimental Conditions.* (i) At least two competent witnesses should always be present at any crucial experiment in addition to the agent and the percipient. They should, if possible, be trained observers—psychologists with experience in mental testing, professional conjurors with a practical knowledge of faked telepathy, communication engineers, or experienced investigators of paranormal phenomena, familiar with all the pitfalls.

(ii) The experiments should be so planned in advance that all known forms of perception, communication, or inference, which would otherwise be likely to affect the results, should be rigorously excluded. This, as we have seen, requires highly specialized knowledge which the psychologist as such is not likely to possess. But I do not see why, with the aid of various technical experts, it should not eventually be possible to prescribe routine procedures and standard devices which, with a little experience and training, the ordinary research worker would be able to adopt.

(iii) Some system of automatic checks is necessary to prevent or eliminate errors in recording the performances of both the agent and the percipient.

(iv) Subject to the foregoing conditions, the general conduct of the sitting should be as natural and informal as possible. As in other types of mental testing, special efforts should be made to put the testees at their ease, to maintain their interest, and to ensure adequate motivation.

(*b*) *Statistical Requirements*[1]. (i) Efficient randomizing procedures should be employed, e.g. a table of random numbers rather than the mere shuffling of cards, and best of all perhaps some kind of mechanical or electronic generator, such as 'Ernie'.

(ii) The point at which the experiments are to be stopped must be decided before the investigation begins.

(iii) The way in which certain parts of the experiments or the data are selected or rejected for the final statistical analysis must not impair the validity of the subsequent statistical inferences.

[1] I have here assumed that the main problem is still to test statistical significance: for the study of epistemological hypotheses the methods of information theory would be more appropriate.

(iv) The type of statistical analysis must be suited both to the hypothesis at issue and to the conditions under which the experiments are performed. This in turn means that these conditions and the general plan of the experiments must be determined in advance with the proposed type of analysis explicitly in view. If a test of significance is required, it will be necessary to choose carefully between a contingency-coefficient, a chi-squared technique, and an ordinary standard-error test based either on the normal curve or the binomial formula or on the postulates of the matching method. For more general types of problem either an analysis of variance, a factor-analysis, or one of the various procedures developed for studying communication and information will be found more appropriate.

APPENDIX

A Note on Statistical Problems

In their discussions of extra-sensory perception many psychologists apparently find it difficult to assess the statistical issues raised. In particular, when told by competent statisticians that in theory the formulae used by Rhine and others are both inaccurate and inappropriate, they do not know how serious are the errors thus introduced. The real need is to find some simple yet valid procedure for estimating the statistical significance of the experimental results. In Dr. Soal's investigations with Glyn something like 15 000 guesses were made in the course of about 600 runs of 25. Obviously it would not do to seize on the most remarkable runs of all, e.g. the two in which every guess was correct, unless we also keep clearly in mind that they were only two out of 600. The simplest way to deal with the problem is to devise an abstract mathematical model embodying the essential conditions, and so deduce formulae and tables giving the theoretical frequencies needed. Three types of model may be suggested according to the conditions of the experiment and the degree of accuracy desired.

(1) *The Normal Distribution.* First of all, let us suppose that, before the experiments on telepathy are carried out, a strictly random sequence of several thousand cards has been compiled by the aid of the tables of random numbers drawn up by Tippett or by Kendall and Babington Smith: (this was in fact the procedure adopted by Dr. Soal in 1934, when checking Rhine's investigations). Then the chance of a correct guess for any card (p say, i.e. 1 out of 5 in Soal's experiments) will remain constant throughout each sitting. In terms of the familiar bag-and-ball model, therefore, the general problem may be re-stated as follows. There are $n = 25$ bags, each containing $s = 5$ balls, 4 black and 1 white: a single ball is drawn at random from each bag; and the whole experiment is repeated $N = 600$ times, making $m = n \times N = 15\,000$ draws in all. In how many cases should we expect to get 0, 1, 2, ... , 25 white balls ?

The theoretical distribution is given by the successive terms of the binomial expansion of the expression $N(p+q)^n$ where $p = 1/s$, and $q = (s-1)/s = 1-p$. It is then commonly argued that, provided p does not greatly differ from q and provided n is not too small, the form of the distribution will be nearly symmetrical, and will approximate to that of the normal curve. Accordingly, most investigators (Rhine, for instance) have used the ordinary table of the normal probability integral, chiefly, it would seem, because such tables are readily accessible and require no special computation.

In the present case we are concerned with the expression $N(\frac{1}{5}+\frac{4}{5})^{25}$. The mean of the distribution will accordingly be $pn \times N = 5N$: in other words, we should expect 20 per cent of the balls actually drawn to be white. The standard deviation will be $\sqrt{\{p(1-p)nN\}} = \sqrt{\{\frac{1}{5} \times \frac{4}{5} \times 25 \times N\}} = 2\sqrt{N}$. Assuming the distribution to be strictly normal, the theoretical probabilities will be those shown in column 1 of Table I.

Suppose, for instance, that in some particular run our subject (Glyn or whoever it may be) draws 9 white balls, instead of 5 as we might expect: can we still reasonably maintain that this large number is just an accidental variation due to the fluctuations of sampling, or should we conclude that some form of trickery, or possibly of extra-sensory perception, is now at work ? The deviation from the mean is $9-5 = 4$. Dividing this by the standard deviation (2 if we put $N = 1$), we obtain for the 'critical ratio'[1] a value of $4/2 = 2$. In other words, the observed deviation is twice the 'standard error'; and, if the distribution is normal, then, as we know a deviation as large as this and in either direction would occur about once in every 20 runs (or slightly more). Here, however, what we want to know, is the probability of a deviation in the *positive* direction; and, with a symmetrical distribution, the odds against this are twice as large—over 40 to 1.

(2) *The Binomial Distribution.* If, however, we are to face statistical criticism, particularly in borderline cases, this rough and ready method is hardly satisfactory. Since p is not in fact equal to q, the actual distribution will be neither strictly normal nor yet perfectly symmetrical; and, since n is a comparatively small number, the 'curve' will be far from smooth. Accordingly, in Table I, column 2, I give the detailed figures for the more exact binomial distribution. It will at once be noted that, up to the level of 9 or 10 correct guesses, the errors involved in following the simpler and more familiar procedure, i.e. using the normal distribution as a convenient approximation, will not be very great. Evidently with either distribution we might expect to find that, in the course of 100 runs, the number of correct guesses with entirely random calls would vary from 0 to 10, and, during 1000 runs, from 0 to 12—a range of just over 6 times the standard deviation: 13 or 14 correct guesses would be decidedly surprising. Thus, up to about this level, which is the highest commonly reached with most

[1] This is what Rhine terms the 'Anti-Chance Index (X)'; in his earlier publications he used the probable error rather than the standard error as divisor.

of the adult percipients hitherto tested, the rough and ready method just described would not after all lead us very far astray.

TABLE I. NORMAL AND BINOMIAL DISTRIBUTIONS

To indicate the probability of correctly guessing 0, 1, 2, ... 25 cards out of a pack of 25, the probability of one correct guess being 1/5.

Correct guesses	Distribution assumed A. Normal	B. Binomial	C. Matching method
0	0·009 245	0·003 778	0·004 286
1	0·027 835	0·023 612	0·025 452
2	0·065 591	0·070 835	0·073 363
3	0·120 977	0·135 767	0·136 572
4	0·174 666	0·186 681	0·184 323
5	0·197 412	0·196 015	0·191 943
6	0·174 666	0·163 346	0·160 257
7	0·120 977	0·110 841	0·110 066
8	0·065 591	0·062 348	0·063 311
9	0·027 835	0·029 442	0·030 896
10	0·009 245	0·011 777	0·012 913
11	0·002 403	0·004 015	0·004 654
12	0·000 488	0·001 171	0·001 453
13	0·000 078	0·000 293	0·000 394
14	0·000 010	0·000 063	0·000 093
15	0·000 001	0·000 012	0·000 019
16	0·000 000*	0·000 002†	0·000 003‡
17	—	238×10^{-9}	531×10^{-9}
18	—	26427×10^{-12}	71150×10^{-12}
19	—	2434×10^{-12}	8247×10^{-12}
20	—	182×10^{-12}	798×10^{-12}
21	—	10866×10^{-15}	70790×10^{-15}
22	—	494×10^{-15}	4010×10^{-15}
23	—	16106×10^{-18}	401×10^{-15}
24	—	335×10^{-18}	0
25	—	3×10^{-18}	16×10^{-18}

* When the number of correct guesses is exceptionally high, then, for reasons stated in the text, the use of a normal distribution is wholly inappropriate.

† Carried to nine decimal places the precise fraction is 0·000 001 797. In the notation used for the remaining probabilities therefore it would be written 1797×10^{-9}.

‡ Carried to nine decimal places the precise fraction is 0·000 003 423.

But Glyn's *average* number was over 10. And many of his *individual* scores ranged from 15 up to 25. Using the normal distribution in the usual way we should infer that a score of 15 or more (i.e. 5 times the standard deviation or over) would occur by chance only once in 3 000 000 runs[1]. But, if we allow

[1] Rhine and other investigators have complained that they " have not found tables for large values of X " (the critical ratio). In *Tables for Statisticians and Biometricians* (Table IV, p. 11) Pearson gives an ' Extension of the Table of the Probability Integral ' running from $X = 5$ to $X = 50$ and thence at wide intervals to $X = 500$.

for the continuity of the normal distribution, we see that the proportion ought rather to be stated as about once in 1 000 000 runs. Using the binomial distribution, however, we should obtain a proportion nearer 1 in 75 000. Nevertheless, when the odds run into thousands, the precise value is seldom of importance. In this particular instance our final conclusion would be the same with either procedure, namely, that the hypothesis of a chance coincidence is quite incredible. Here, however, my point is rather to illustrate the possible risks: in other investigations of a similar type the use of the more familiar method in place of a binomial distribution might expose the investigator to serious criticism.

To show how extraordinary the results really are let us consider the probability of obtaining a result of 20 or more. From the figures for the binomial distribution we see that by chance such an achievement would occur about once in 5000 million runs. Now it takes a quick worker between 3 and 4 minutes to run through a pack of 25 cards. Thus, working continuously both night and day during the whole of his life, Glyn would only have managed about 2 million runs. To accomplish 5000 million he would have to have been guessing from the time of the palaeolithic man. Actually, however, he achieved this result 26 times in less than 10 000 trials, i.e. 400 runs, or what would be little more than a single day of continuous working.

(3) *The Matching Method.* Now in most of Rhine's experiments, and in those carried out by Dr. Soal with Glyn and Ieuan, the packs of cards were not compiled in the way described above. In each pack and in every run there were *exactly* five cards of each of the five possible types (geometrical designs in Rhine's experiments animals in most of Soal's). In this case we can no longer assume that the chances of a correct and an incorrect guess remain the same throughout the run. For, if the first card drawn is of type Z(e.g. a Zebra), then, unless it is replaced in the pack, the chance that the next will be of type Z is no longer 5/25 but 4/24; or again suppose (as will occasionally happen in a long series) that the first five cards are all of type Z, then at the sixth guess the chance that the call 'Zebra' will be correct has already dropped to zero, since no more Zebras are left in the pack: (with a *random* compilation of cards, i.e. with a pack which itself had been based on random draws, the probability would of course in both cases still be 5/25). Finally, suppose that the percipient knows that there are five cards of each type and arranges his guesses accordingly, and suppose that he has correctly guessed 24 cards (though he may not know it): then his last guess is *bound* to be correct; indeed, when the guesses include five of each type, then, unless the percipient guesses *every* card correctly, his highest possible score will be only 23 not 24 (see Table I, col. 3).

Thus, with fixed numbers for each type that the pack contains, the problem takes a very different form: it has in fact become a problem in *matching*—a point which many experimenters in telepathy have apparently overlooked. Accordingly, to start with a mathematical model of a fairly general form, we may describe the situation in the following terms. We have first a 'target'

pack of n cards, which we can imagine laid out in the order in which the 'agent' transmits their names; they comprise s different types or 'suits', e.g. a_1 Lions, a_2 Elephants, ... a_s Zebras. Next, we may imagine that a second pack of n cards is laid out one by one beneath the first row, to represent the percipient's guesses: they will, as before, consist of the same s types[1], but the number of times each type is called may be different. Let us suppose that there are b_1 guesses of the first type, b_2 of the second, ... b_s of the last: then

$$a_1+a_2+ \ldots +a_s = b_1+b_2+ \ldots +b_s = n. \qquad \text{(i)}$$

Lastly, let r designate the number of cases in which the guessed card *matches* the target card: in other words, suppose that the number of 'correct' guesses is r out of n. Our problem is to find the general probability-distribution of r.

Stated in this generalized fashion, we have a matching problem similar to that which frequently crops up in other psychological inquiries[2]. The s 'target types', for instance, might be a set of temperamental or occupational 'types': and the 'guesses' a set of descriptive assessments for such types drawn up by an assessor or judge. Since the procedure lends itself to so many different kinds of investigation, it will be worth while to examine it at somewhat greater length.

Let us begin with the simplest case of all, that in which

$$a_1= \ldots = a_s = b_1= \ldots = b_s = 1,$$

i.e. that in which *every* card is of a different type so that $s = n$: this will make the basic method of solution clear; and the reasoning can then readily be generalized to the more complicated cases. Suppose that the n cards are labelled $A, B, C, \ldots, N$. They are shuffled, and dealt out in random order, underneath a similar series of 'target cards'. No generality will be lost if we assume the target cards to be laid out in a standard order, e.g. alphabetical. We now want to know what is the probability that 1, 2, 3, ... , or n cards among those thus dealt out will be found in their correct alphabetical position[2].

Under the name of the 'game of encounters' (*rencontres*) a gamble of this kind was familiar to early French card-players; and a solution to the problem was first given by P. de Montmort (a pupil of Malebranche) in his *Essai d'analyse sur les jeux de hasard* (Paris, 1708; 2nd ed., 1714, pp. 140–4: this edition contains a generalization by N. Bernoulli). A. de Moivre's discussion (*Doctrine of Chances*, 1718, problem XXXV) is rather fuller, and brings out an important

[1] If there are any types in the target series which do not appear in the series of guesses, we can designate their number by a_0. Similarly, if the 'guesses' include answers that correspond to no cards in the target series (e.g. such answers as 'I don't know' or 'Nothing comes to my mind'), we can represent these by blank cards, and designate their number by b_0.

[2] An alternative scheme is to set out the results in the form of a square $s \times s$ contingency table: (cf. W. L. Stevens, op. cit. *inf.*) Contingency coefficients calculated by Pearson's procedure have not infrequently been used in psychological studies of personality based on a matching method: (cf. D. W. Chapman, *Amer. J. Psychol.* v, 1934, pp. 37–54 and refs.). But it is clear that such a coefficient answers a somewhat different question: it is in fact affected by the way the correct guesses are distributed among the different types of card. On the other hand, this approach would offer a simple procedure for measuring the *strength* of the E.S.P. capacity of given persons (or of given situations): (cf. C. Burt, 'The Factorial Analysis of Qualitative Data' this *Journal*, III, p. 177 f.).

principle which applies equally well to the more complicated cases—namely, the principle of successive 'inclusion and exclusion'. He uses $+A$ to denote the chance that A is in its right place, and $-A$ the chance that it is not; and similarly for the other cards. Then the chance that A is in its right place $(+A)$ will be $(n-1)!/n! = 1/n$; the chance that both A and B are in their right places $(+A+B)$ will be $(n-2)!/n! = 1/[n(n-1)]$; and hence the chance that A yields a match and B does not $(+A-B)$ will be $1/n - 1/[n(n-1)]$. De Moivre sets out the results progressively obtained by following this simple type of reasoning according to the scheme shown below: however, it will be more convenient to substitute for his symbols the modern notation for factorial powers, viz. $n^{-(1)} = 1/n = (n-1)!/n!$, $n^{-(2)} = 1/[n(n-1)] = (n-2)!/n!$, and so on: then

$+B$	$= n^{-(1)}$
$+B+A$	$= n^{-(2)}$
$+B-A$	$= n^{-(1)} - n^{-(2)}$
$+C+B$	$= n^{-(2)}$
$+C+B+A$	$= n^{-(3)}$
$+C+B-A$	$= n^{-(2)} - n^{-(3)}$
$+C-A$	$= n^{-(1)} - n^{-(2)}$
$+C-A+B$	$= n^{-(2)} - n^{-(3)}$
$+C-A-B$	$= n^{-(1)} - 2n^{-(2)} + n^{-(3)}$
$+D-B-A-C$	$= n^{-(1)} - 3n^{-(2)} + 3n^{-(3)} - n^{-(4)}$

It will be seen that the values obtained when only one card is in its right place and 1, 2, 3, ... , j cards are out of place are analogous to the numerical coefficients obtained on multiplying $(1-n)$ by itself j times; i.e. the proportional frequencies are the binomial coefficients of the successive terms in the expansion of $(1-n)^j$. The alternation of $+$ and $-$ signs in the results on the right-hand side shows how we begin by *including* a certain class of cases, and then *exclude* those not required, and then *include* those wrongly excluded, and so on[1].

[1] The non-mathematical student will follow the argument more readily if he puts $n = 4$, and sets out all the $4.3.2.1 = 24$ possible arrangements of the four letters, and then verifies the result given by the last line of the table above: viz., that the probability that one particular letter, say D, shall be in its right place is $2/24 = 1/12$, and that consequently the probability that any one of the four shall be in its right place is $4 \times 1/12 = 1/3$. He should then derive and verify the following more general formula, viz., the probability that r letters out of n shall be correctly placed is

$$P_n(r) = [1 - 1/1! + 1/2! - \ldots + (-1)^{n-r}/(n-r)!]/r! \ldots . \qquad \text{(ii)}$$

It is instructive to note that the series in square brackets, if continued indefinitely, is the value of e^{-1}, where e denotes the base of the natural logarithms (2·718 ...). This yields a rough and ready way of assessing the several probabilities.

Let us now turn to the more general case in which the composition of the two packs is that indicated by equation (i). As before, it is necessary to suppose that the order of the cards in *one* of the two series is random. In experiments on E.S.P. it is commonly assumed that, as a result of thorough shuffling or (better) of using a table of random numbers, the order in the *target* is random. The other set may then be in a fixed or any other order: e.g. the 'percipient' may, if he likes, use sequences of guesses—e.g. *AAAAA* (all 'lions' say), then *BBBBB*, and so on[1].

Various methods of proving the appropriate formula are available in the relevant literature: the general mode of reasoning is the same as that described above for the simpler case; and it is unnecessary to elaborate it here[2]. The formula itself can also be expressed in several different ways. For purposes of practical computation the following is perhaps the most convenient. If, as before, we denote the probability of r correct matches by $P(r)$, then

$$P(r) = \frac{1}{n!} \sum_{j=r}^{n} (-1)^{j-r} \frac{j!}{r!(j-r)!} (n-j)! C_j, \qquad \text{(iii)}$$

where C_j is the coefficient of x in the expansion of

$$\prod_{i=0} \left[\sum_{k=0} \frac{a_i! b_i! x^k}{(a_i-k)!(b_i-k)!k!} \right], \qquad \text{(iv)}$$

[1] The model suggested thus serves for a *conditional* distribution, i.e. one in which the *relation* between the target card and the guessed card is random. But, let us note, the assumptions involved will scarcely be satisfied if the *same* target pack is just shuffled and re-shuffled for each new run, since the previous experience of the guesser might lead him to hit on a sequence which *was* in some degree related to the sequence of cards in the pack before the shuffling; and (as has often been shown) such sequences are not *completely* deranged by shuffling and cutting in the ordinary manner. In any case the common statement that, if the guesser is not guided by some normal or para-normal source of information, then *his* guesses will be random, is quite erroneous: when guessing blindly, most subjects tend to space out the different types of card more or less uniformly, so that runs of the same type (which are bound to crop up from time to time with a genuinely random distribution) are thus excluded. In my view, therefore, the experimenter should always use random numbers to decide the order of the target cards.

One consequence of the assumptions on which the formulae are based seems to have been generally overlooked. As will be seen, the question we now have to ask is not the usual one—viz "given the order of the target cards, how many of the guesses are likely to correspond, if chance alone is operative?" but "given the order of the cards guessed, how many of the target cards (which we assume to be arranged in random order) are likely to correspond with the guesses?" With a *single* percipient only, the answers are the same in both cases. But where, as in many *group* experiments, all the percipients guess at the *same* target series, we can no longer add their results and apply the usual formulae to the totals. The reason is that we no longer have grounds for assuming that the guesses are independent. Indeed, as experience shows, the more familiar animals (or, with ordinary packs, the court cards) tend to be guessed first.

[2] For the psychologist the most useful discussions are those of W. L. Stevens ('The Distribution of Entries in a Contingency Table', *Ann. Eugen.*, VIII, 1939, pp. 238–244), I. L. Battin ('On the Problem of Multiple Matching', *Ann. Math. Statist.*, XIII, 1942, pp. 294–305), E. J. Gilbert ('The Matching Problem', *Psychometrika*, XXI, 1956, pp. 253–266), and D. E. Barton ('The Matching Distributions', *J. Roy. Statist. Soc.*, B. XX, 1958, pp. 73–92). An instructive account of the subject in its general mathematical context will be found in J. Riordan, *An Introduction to Combinatorial Analysis* (1958), esp. chaps. III and VII.

the summation being taken from 0 to a_i or b_i, whichever is the smaller. The mean of the distribution, which gives the number of matches to be expected by chance, and the standard deviation are given by the following equations:

$$\bar{r} = \frac{1}{n}\sum a_i b_i, \tag{v}$$

$$\sigma_r^2 = \frac{1}{n^2(n-1)}\left[(\sum a_i b_i)^2 - n\sum a_i b_i(a_i+b_i) + n^2\sum a_i b_i\right]. \tag{vi}$$

Now, in the type of experiment with which we are here concerned, the number of cards of each type or suit in the target is the same. Hence $\sum a_i b_i = a\sum b_i = an$. Consequently

$$\bar{r} = a, \tag{vii}$$

and

$$\sigma_r^2 = \frac{a}{n(n-1)}\left[n^2 - \sum b_i^2\right]. \tag{viii}$$

But Σb_i^2 is a minimum when $b_1 = b_2 = \ldots = b_k$. Therefore, when the number of each type in the series guessed is the same (viz., $n/s = a$), the variance is a maximum, and then has the value

$$\sigma_r^2 = \frac{a}{n(n-1)}\left[n^2 - sa^2\right] = \frac{n^2}{s^2}\cdot\frac{s-1}{n-1}. \tag{ix}$$

We may write this

$$\sigma_r^2 = \frac{n^2}{n-1}\cdot\frac{1}{s}\cdot\frac{s-1}{s} = \frac{n}{n-1}\cdot npq, = \frac{n}{n-1}\sigma^2_{\text{bin}},$$

where σ^2_{bin} denotes the variance of the corresponding binomial distribution. Hence, the effect of compiling the target pack by taking the same number of each type instead of selecting the cards at random is to increase the standard deviation by a factor of $\sqrt{\{n/(n-1)\}}$, a factor which approaches 1 as n becomes large. Thus, in considering the significance of any specific experimental result, we can often simplify the necessary calculations, by taking the maximum variance. Accordingly, if (as in the experiments discussed in the text) $n = 25$, $s = 5$, and $N =$ the total number of trials, the standard deviation becomes

$$\sqrt{\left\{\frac{25}{5^2}\cdot\frac{4}{24}\right\}}\times\sqrt{N} = \sqrt{\frac{1}{6}}\times\sqrt{N},$$

that is $0{\cdot}408\sqrt{N}$, instead of $0{\cdot}400\sqrt{N}$.

Similarly the generating function given above (iv) for calculating the coefficients of x_j reduces to

$$\left[\sum_{k=0}^{a}\frac{(a!)^2x^k}{\{(a-k)!\}^2k!}\right]^s. \tag{x}$$

When s is a fairly small number (as is the case in most of the matching problems with which the psychologist is usually concerned), the calculations are not very

laborious[1]. The first step is to calculate the sum in formula x (or the sums in formula iv) and then expand the power (or the product) in terms of ascending powers of x. The factorials and the binomial coefficients can readily be found from mathematical tables[1]. For the present series of experiments the values thus obtained are shown in column 3 of Table I.

The relations between the various types of distribution to which we have referred may be illustrated by comparing the main constants for each—the mean and the second and third moments about the mean. These are indicated in the following table:

	Distribution			
Constant	Normal	Binomial	Matching method	Poisson
Mean μ_1	np	np	$np=a$	np
2nd Moment μ_2	npq	npq	$npq[n/(n-1)]$	np
3rd Moment μ_3	0	$npq(q-p)$	$npq(q-p)[n^2/(n-1)(n-2)]$	np

As n increases in size so the fractions in square brackets approach unity; and therefore both the second moment (the variance) and the third moment for the matching method approach those for the binomial distribution. When the proportions p and q become equal, the third moment of the binomial vanishes, and the distribution becomes symmetrical. On the other hand, when p becomes very small (and q correspondingly large), q approaches unity; and the second and third moments for the binomial distribution and the matching method approach those of the Poisson distribution.

When, as in the case of the present experiments, $p = 1/5$ and $n = 25$, and when the number of cards in type or suit are fixed in the way described above, then (as is shown by the figures for the matching method in Table I) the probability of obtaining a large number of matches, though still extremely small, is appreciably higher than we should infer by using the ordinary binomial distribution. For Glyn's guesses the matching method seems unquestionably the most appropriate; and for a score of 15 or more this gives odds of 1 in 41,500 only. In general, however, except in borderline cases, the difference is hardly large enough to alter the conclusions drawn. Thus, as Fisher has observed, the real task of psychical research workers is, not so much to refine their statistical procedures, but rather to improve their experimental techniques; and this is precisely what Dr. Soal has endeavoured to do.

[1] By way of exercise the student is recommended to apply both formulae to the case in which there are only 6 cards and only 3 different letters or 'suits', taking $a = n/s = 2$ for formula (x). He should also satisfy himself that the fuller formulae reduce to equation (ii) for the case in which $a = 1$ (the *problème des recontres*).

In practice, when s is large, instead of calculating the exact values by using the formulae given above, we can substitute as an approximation some type of continuous curve: usually the binomial gives a better fit than the normal, and a still better fit may be obtained by substituting a modified Poisson distribution, a Pearson Type I, or a Gram-Charlier Type B. See Riordan and Barton, *loc. cit. sup.*

Postscript. The criticisms advanced by Mr. Hansel and discussed on p. 59 f. relate to Dr. Soal's earlier work. At that time Mr. Hansel believed that " possibly never again would a percipient be produced who can consistently obtain high scores ". The Jones boys, however, have now produced higher scores than any previous percipient. Accordingly, as in this case his earlier explanation was ruled out by the procedure adopted, he has since put forward in a more recent note (*New Scientist*, 30 April, 1930) two new suggestions to account for their unexpected achievements. These have been widely accepted by psychologists and others who have not read Dr. Soal's book, but have assumed that in all essentials Mr. Hansel's summary is accurate and complete. This, however, is far from the case.

(i) First, he states that " under condition 2 " (i.e. when the experiments were carried out indoors with the boys at a distance from each other) " the Jones boys ceased to obtain high scores when the agent's legs were not visible to the percipient ". Now Dr. Soal expressly emphasizes that, when the boys were brought to London and the experiments were conducted in the large rooms of the S.P.R., " Glyn was unable to see *any* part of Ieuan's body *including his feet and legs* " (p. 59). No member of the family was present. Yet extremely high scores were obtained: in a series of 12 runs (excluding the first two when the parents were present) scores of 13 or over were obtained during seven of the runs, and scores of 18 during two more. Even had Ieuan's feet been visible, it would scarcely be plausible to suppose (as Mr. Hansel does) that the trained observers who watched the experiments continuously would have failed to detect foot-signals that were visible to Glyn at 24 feet; and it is surely quite unfair for Mr. Hansel to say that he " suspects that visual cues were given by the agent ", without adding that the investigators themselves have stated that the agent was " *completely screened* " (p. 63).

(ii) Mr. Hansel's alternative suggestion is that the boys might have devised a code based on sounds produced by a high-pitched whistle; this, he thinks, could have been concealed in a trouser pocket and blown with a bulb. In an experiment he himself carried out with two children at a distance of 50 ft. a score of 16 out of 25, he says, was achieved by this method. He apparently failed to notice that, in Dr. Soal's experiments, scores of 20 to 23 out of 25 were obtained at a distance of 166 feet. To this Mr. Hansel has since replied that " with practice we could easily have obtained a score of 20 out of 25 at a distance of 150 ft. " If it was as easy as all that, why did he not actually repeat the experiments at this distance instead of asking us to accept his assurance that it *could* be done ? I have myself carried out experiments with a similar whistle on a number of children of school age; and I find that, even with those whose hearing is exceptionally acute, it is impossible to make a high-pitched sound audible to the child at 160 ft. or more, while it remains inaudible to young adults (like some of Dr. Soal's observers) standing only 2 or 3 ft. away. Moreover Mr. Hansel overlooks the highly successful experiments which were carried out at the distance mentioned when the boys wore close-fitting bathing trunks and were carefully searched. No other member of the Jones family was present.

However, at the very outset of his published comments Mr. Hansel was evidently so confident that it is now practicable to state in advance what results are or are not scientifically possible in experiments of this kind that, as he frankly observes, he starts by " assuming that telepathy and precognition are impossible ": " we *know* ", he adds, " that there must be some explanation of the trick which involves only natural processes "—telepathic communication being in his view essentially unnatural. Armed with this antecedent knowledge, therefore, he plainly deems it waste of time to seek empirical evidence to confirm his imaginary explanations. It need hardly be said that this armchair method of dismissing awkward observations *a priori* is not likely to contribute greatly to the advance of scientific research.

BOOK REVIEWS

Eleven-Year-Olds Grow Up. By J. S. MACPHERSON. University of London Press, Ltd., 1958. Pp. xxiv+186. 15*s*.

In their report on *The Trend of Scottish Intelligence* (1959) the Scottish Council for Research in Education announced that their Mental Survey Committee had in view a plan for following up, over a period of at least twenty years, a large and representative sample of the pupils tested in the original inquiry. Thanks to the generosity of the Nuffield and the Rockefeller Foundations and to the assistance given by a keen band of voluntary workers, it has now been possible to trace the history of a number of these children—namely, the 1,208 boys and girls who formed the ' six-day sample '—during the first seven years that followed the survey. The detailed results, which relate chiefly to the educational, vocational, and social progress of the group, are here presented in the form of some eighty-six tables with a running commentary by the Council's Research Officer, Mr. Macpherson.

The most important conclusions are summarized in the final chapter. This deals more particularly with the achievements of ' high and low scorers ', i.e. those whose I.Q.'s amounted to 132 or more, or 77 or less. In contrast to the ' lower scorers ', the typical ' high scorer ', we are told, is likely to come from a relatively small family. His father will most probably be a member of the professional class, and in any case is not likely to have an occupation inferior to that of a skilled tradesman. By the age of 11, the boy will probably be selected for a five-year secondary school course, in which he will do well, and, if we may trust the teachers' reports, be ' very happy '. His school attendance will be excellent—at least 95 per cent of what is possible; and his masters will usually rate him as above the average in the more desirable qualities of personality. He is likely to spend at least five years in the secondary school, and on leaving will have a fifty–fifty chance of starting to train for a profession.

A small proportion of the high scorers, however, failed to come up to the expected level either educationally or vocationally. In these cases, it appears, the commonest reason is some deficiency of character: the child is described as ' lazy ', ' indifferent ', ' lacking in persistence ', or ' disliking hard work ' ; in a few instances, however, ill health may be to blame. Mr. Macpherson concludes that " there is a substantial relation between a pupil's score in the intelligence test at age eleven and his subsequent progress ", but that " social factors also play a considerable part ".

As he words them, the conclusions sound a little trite and obvious, and might perhaps have been expressed with greater precision. It would too have helped busy readers of the report had he attempted some kind of statistical reduction so as to condense the tabulated data to terms of a few pregnant figures. However, we are told that the present report forms only " the first stage of the follow-up ". The details given are full of interest; and we may hope that, when the time comes, the final survey will be still more illuminating.

E. F. WATSON.

Statistics in Psychology & Education. By H. E. GARRETT. London: Longmans, Green & Co. Pp. xiv+460.

This is the fourth edition of Dr. Garrett's well-known textbook. As he states in the new preface, in spite of the enlargements his purpose remains the same as when the book was first published, namely " to present those fundamental principles of statistical method which will be most useful to students in psychology and education ". It is thus essentially an elementary course for beginners, and highly specialized topics are not included. The chapters on reliability and inference have been completely rewritten; a new chapter has been added dealing with the analysis of variance; several obsolescent techniques have been omitted; and the whole has been completely revised and reset. R. S. JOHNSON.

Thinking: An Experimental and Social Study. By Sir Frederic Bartlett. London: George Allen & Unwin, 1958. Pp. 204. 18*s*.

Professor Bartlett tells us he has had three principal aims in view: (i) " to put thinking in its place as a natural development from earlier forms of bodily skilled behaviour "; (ii) on this basis, " to design more or less novel experiments such that anybody with sufficient interest could carry them further for himself "; (iii) " to illustrate the experiments and embark on some discussion of the results ". ' Thinking ' he classifies into four main types. He begins with ' thinking within a closed system ' as the form most amenable to experimental control; and then goes on to discuss the main types of ' adventurous thinking ' (as he terms it), namely, ' everyday thinking ', ' experimental thinking ', and 'the artist's thinking'. A valuable feature of his experiments is the free use made of introspective reports from those who served as subjects: the results fully justify the value of this old-fashioned procedure, which, owing to the strictures of the behaviourists, has dropped into undeserved disfavour.

The guiding principle that animates his work, however, is a firm conviction that thinking is not only an outgrowth from more primitive types of skill, but is itself to be interpreted by analogy with these lowlier forms of behaviour. At times he seems to push the parallel a little too far: he argues, for example, that, just as " the most important characteristic of expert bodily skill is timing ", so also in everyday thinking " successful practitioners are especially sensitive to the timing ". It would seem to be the differences rather than the similarities between the two that really call for detailed analysis.

Since the book is intended primarily to throw out suggestions for others to follow, it would be unfair to emphasize its obvious limitations and omissions. However, unless attention is expressly drawn to them, there is a danger that research students who seek to carry out the author's programme may be tempted to adopt his rather simplified techniques. Professor Bartletts inferences are presented in an easy-going colloquial style which undoubtedly is far pleasanter to read than (say) Miss Wheeler's dry statistical treatment of the subject in earlier issues of this journal. Nevertheless her arguments are far more convincing; and, were it not for the shrewd common sense and prolonged experience which manifestly underlies Professor Bartlett's contentions, a critical reader might be tempted to question how far the evidence which he offers is really conclusive. In one place he assures us that, " so far as my results go, about two people out of every hundred taken at random may be expected to produce " a given type of reply; but we are left in the dark as to how many persons were in fact tested, or what methods (if any) were adopted to ensure that they really formed a random sample. A little later he says that one mode of solution is, " according to my results, just about twice as likely as the latter "; but no statistical data are given to support this estimate. On the whole it seems highly likely that many of the conclusions reached may hold good only of the particular subjects who took part in the experiments—chiefly, it would appear, Cambridge graduates or undergraduates.

A second peculiarity is the omission of almost all references to the methods used and the hypotheses advanced by previous investigators. After all, from the days of Binet and the Würzburg school, a vast amount of research has been carried out on these problems. At one stage of his work Professor Bartlett " began to wonder," so he tells us, " whether there might be a direct relation between capacity to utilize minimal information so as to build up sequential arrangements and high ranking intelligence "; and he argues the question out in general terms. But a definite answer was given long ago when Burt showed that differences in intelligence could best be estimated by the complexity of the most difficult type of reasoning-problem which each examinee was able to solve[1]: many of the puzzles used in

[1] Burt, C., Experimental Tests of Higher Mental Processes and their Relation to Intelligence, *J. Exp. Pedag.*, I, 1911, pp. 93–112. The parallelism was most obvious in the ' symbol substitution tests ', where a coded message or multiplication sum in letters has to be solved, and the examinee is deliberately aided by having 1, 2, 3 or more symbols decoded for him in advance. Both types of problem are used with excellent effect by Bartlett.

Professor Bartlett's experiments seem in fact to have been suggested by the problems in reasoning which were regularly used in these early group tests of intellience. His views on the ' extraction of rules ' and the ' application of rules ' bear a close resemblance to those put forward by Stout and Spearman in explaining what the latter calls the ' eduction of relations and of correlates '. But it is curious that, unlike other investigators, he has apparently encountered no evidence for the important rôle played by what Stout called ' schemata ', a concept which figured conspicuously in Bartlett's own work on ' remembering '. Finally, in his account of the way in which thinking develops out of more primitive types of problem-solving, it is strange that no mention is made of the recent researches of Piaget and his collaborators.

These, however, are minor defects. As Professor Bartlett points out, his discussion was intended to be " in no sense final or systematic ", but merely to sketch lines of approach which others might follow up. And unquestionably his book is one which every psychologist interested in the development of human reasoning, and every investigator engaged in research on thought processes, should feel bound to study and ponder. R. B. LEWIS.

Science, Theory, and Man. By E. SCHROEDINGER. London : George Allen & Unwin, 1957. Pp. 223. 18*s*.

This is a new edition of a series of lectures by Professor Schroedinger published some years ago under the title of " Science and the Human Temperament ", to which a recent essay on ' the nature of an elementary particle ' has been added. The volume includes a biography of Professor Schroedinger by Dr. James Murphy, and a preface to the original edition by Lord Rutherford. As Dr. Murphy observes, the book itself is " of the nature of a pilgrimage to the Cathedral of Cloyne "; for Schroedinger, like Jeans, Eddington, and Whitehead, finds that modern physics " leads away from the materialism of the behaviourists to the mentalism of which Bishop Berkeley was so eloquent an exponent ". In view of the recent discussions in this journal of the bearings of quantum theory and wave mechanics on the work of the psychologist, Professor Schroedinger's lucid and authoritative examination of these problems should be of special interest; and those who have taken part in the current debate on ' scientific methodology ' might do well to study his chapters on indeterminism, the laws of chance, the notion of a law of nature, and the philosophic value of conceptual models in scientific inquiry. His own conclusion is that even the most exact and objective of the sciences are " dependent on the frame of mind fashionable during the epoch of which it forms a part ", or, as he puts it (modifying an apothegm of Zola) *la science c'est la nature vue au travers d'un tempérament.* W. L. GREGORY.

This is my Philosophy. Edited by WHIT BURNETT. London: George Allen & Unwin, 1958. Pp. xxii+378. 25*s*.

It has been said that, " unlike their predecessors, psychologists of today approach human problems lamentably ignorant of current philosophic trends, and the statistical psychologist has the narrowest outlook of all: he sees nothing beyond his blinkers ". If so, this is a volume which should help to enlighten him. It consists of excerpts from " twenty outstanding thinkers of the present day ", intended to present (so far as is possible in the space of the few pages allocated to each) the essence of their " philosophical message ". A few of the contributors are, like Bertrand Russell, Professor Hocking, and Professor Jaspers philosophers in the academic sense. Some, like Sartre, Maritain, Marcel, and Jung, have had great influence on current academic thought. The rest represent a number of widely different fields: they include physicists like Heisenberg, biologists like J. B. S. Haldane, historians like G. M. Trevelyan, novelists like Aldous Huxley, and an architect—Frank Lloyd Wright. All the discussions, however, are well worth studying; and, together with the biographical introductions, would do much to broaden the outlook of the young student of psychology. J. F. MASON.

Television and the Child. By H. T. Himmelweit, A. N. Oppenheim, and P. Vance. London: Oxford University Press, 1958. Pp. xxii + 522. 42*s*.

The investigations described by Dr. Himmelweit and her colleagues were intended to determine " what are the chief effects of television programmes on viewers of school age ". They are based on studies carried out in 1954–5 with two main age-groups—namely, children aged 10 to 11 in the top classes of primary schools and children aged 13 to 14 from secondary modern and grammar schools. The principal survey was conducted in four large cities (London, Portsmouth, Sunderland, and Bristol) and included a sample of 927 ' viewers' who were matched for age, sex, intelligence, and social background with an equal number of ' controls '. A supplementary survey was carried out in Norwich, and included 185 children who had access to television sets after a new transmitter had been opened and who had previously been tested before television was available. The data thus obtained are summarized in 73 tables, and the conclusions discussed under six main headings.

It appears that, at the time of the inquiry, children in both age groups watched on an average for about two hours every day: they spent more time on television than on any other leisure activity. At least half the children watched programmes for adults, often sitting up after 10 p.m. to do so. In a questionnaire on preferences it was found that " three quarters of the votes for the most favoured programme went to adult programmes—particularly to crime thrillers and to a less extent comedies and variety . . . . The single most important background factor was undoubtedly intelligence: the higher the child's intelligence, the less his viewing ". The writers conclude that the appeal of television arises from the fact that it " provides escape from everyday demands to glamour, excitement, and romance ". The gain in general knowledge was extremely slight; and there was no indication that viewing aroused interest in any particular school subject. Strangely enough, viewing (so we are told) in no way reduced the amount of time spent on homework: however, here the investigators admit that " a repeat inquiry is needed ". On the whole, they conclude, the fears so commonly expressed about the deleterious effects of television—on character, school progress, eyesight, sleep, health and outdoor games—have little basis in fact. Here as elsewhere the truth of the matter would seem to be that the ill effects that can unquestionably be detected in a few sporadic cases are seized on by the casual observer and made the foundation for sweeping generalizations which find little or no support in carefully conducted statistical surveys. " All in all," we are told, " the values of television can make a wholesome impact if they are consistently presented in dramatic form and touch on ideas or values for which the child is emotionally ready ".

It would be easy to criticize both the planning of the investigation and the analysis of the data from the standpoint of ideal statistical requirements. For example, the sample of ten-year-olds was limited solely to children in the top classes, and is therefore hardly representative of the whole population; and the surveys themselves were carried out before two channels became available. Moreover, the results are based mainly on what children and their parents say about their habits and their interests; and it may be questioned whether youngsters of ten in primary schools are able to give clear and objective answers on paper to the type of questions put. The comparison of the viewers with controls is of a rather rough and ready type; and the way the conclusions are worded suggests that often they reflect the investigators' impressions and preconceived theories rather than the quantitative data actually set out in the tables. In the attempts to measure " associations between programme types " Kendall's ' Tau ' was used; but the correlations thus obtained are due in many cases more to an overlapping of actual items than to interest or ability. This is plainly a problem which should be tackled by factor analysis—a method which for some unstated reason the investigators suppose was inapplicable. However, the investigation was essentially a pioneer survey: many of the apparent shortcomings—which the writers themselves frankly acknowledge—were due to practical exigencies. The more important conclusions and the practical recommendations based on them seem sound in themselves, and are clearly stated and substantially supported.

Charlotte Banks.